KAISA TEITTINEN

Nucleolar RNAs and Proteins in Leukemia and Zebrafish Development

ACADEMIC DISSERTATION
To be presented, with the permission of
the Board of the School of Medicine of the University of Tampere,
for public discussion in the Small Auditorium of Building B,
School of Medicine of the University of Tampere,
Medisiinarinkatu 3, Tampere, on February 20th, 2014, at 12 o'clock.

UNIVERSITY OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere, School of Medicine
Tampere University Hospital, Department of Paediatrics
Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB)
Finland

Supervised by
Docent Olli Lohi
University of Tampere
Finland
Professor Markku Mäki
University of Tampere
Finland

Reviewed by
Docent Matti Korhonen
University of Helsinki
Finland
Docent Maija Vihinen-Ranta
University of Jyväskylä
Finland

Cover design by
Mikko Reinikka

Acta Universitatis Tamperensis 1902
ISBN 978-951-44-9361-4 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1384
ISBN 978-951-44-9362-1 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2014

You can't start a fire without a spark.

-Bruce Springsteen

Contents

List of original communications

This thesis is based on the following original communications, which are referred to in the text by their Roman numerals I-IV.

I **Teittinen KJ**, Laiho A, Uusimäki A, Pursiheimo JP, Gyenesei A, Lohi O. Expression of small nucleolar RNAs in leukemic cells. Cell Oncol (Dordr). 2013, 36:55-63.

II Liuksiala T, **Teittinen KJ**, Granberg K, Heinäniemi M, Annala M, Mäki M, Nykter M, Lohi O. Overexpression of SNORD114-3 marks acute promyelocytic leukemia. Leukemia. 2013, doi: 10.1038/leu.2013.250

III **Teittinen KJ**, Kärkkäinen P, Salonen J, Rönnholm G, Korkeamäki H, Vihinen M, Kalkkinen N, Lohi O. Nucleolar proteins with altered expression in leukemic cell lines. Leuk Res. 2012, 36:232-236.

IV **Teittinen KJ**, Grönroos T, Parikka M, Junttila S, Uusimäki A, Laiho A, Korkeamäki H, Kurppa K, Turpeinen H, Pesu M, Gyenesei A, Rämet M, Lohi O. SAP30L (Sin3A-associated protein 30-like) is involved in regulation of cardiac development and hematopoiesis in zebrafish embryos. J Cell Biochem. 2012, 113:3843-3852.

Abbreviations

2-D DIGE	two-dimensional difference gel electrophoresis
2-DE	two-dimensional gel electrophoresis
3' UTR	3' terminal untranslated region
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
AP	alkaline phosphatase
APL	acute promyelocytic leukemia
BLL	Burkitt's lymphoma/leukemia
bp	base pair(s)
BSA	bovine serum albumin
cDNA	complementary DNA
CLL	chronic lymphatic leukemia
CML	chronic myeloid leukemia
CNS	central nervous system
DE	differentially expressed
DFC	dense fibrillar component
DIG	digoxigenin
DKC1	dyskerin
DNA	deoxyribonucleic acid
dpf	days post-fertilization
EF1A	elongation factor 1 alpha
FAB	French-American-British
FBS	fetal bovine serum
FC	1) fibrillar center 2) fold change
GC	granular component
HIV	human immunodeficiency virus
hpf	hours post-fertilization
HRP	horse radish peroxidase
HSC	hematopoietic stem cell
HSPC	hematopoietic stem-progenitor cell

HSV-1	herpes simplex virus 1
ICR	imprinting control region
IEF	isoelectric focusing
IGS	intergenic spacer
IR	ionizing radiation
JAK	janus kinase
kb	kilo base (pairs)
lncRNA	long non-protein coding RNA
MDS	myelodysplastic syndrome
miRNA	microRNA
MLL	mixed lineage leukemia
MNC	mononuclear cell
MO	morpholino
MRD	minimal residual disease
mRNA	messenger-RNA
MS	mass spectrometry
NCL	nucleolin
ncRNA	non-protein-coding RNA
NLS	nuclear localization signal
NoDP	nucleolar detention pathway
NoDS	nucleolar detention signal
NoLS	nucleolar localization signal
NOR	nucleolar organizer region
NPM	nucleophosmin
nt	nucleotide(s)
PBS(T)	phosphate buffered saline (with Tween-20 detergent)
PCR	polymerase chain reaction
PFA	paraformaldehyde
Ph	Philadelphia chromosome
PHB	prohibitin
PNB	prenucleolar body
PP1	protein phosphatase 1
ppox	protoporphyrinogen oxidase
PTCL	peripheral T-cell lymphoma
PTEN	phosphatase and tensin homolog

PWS	Prader-Willi syndrome
RARA	retinoic acid receptor alpha
Rb	retinoblastoma protein
RC	random control (morpholino)
rDNA	ribosomal DNA
RP(L/S)	ribosomal protein (of the Large/Small subunit)
rRNA	ribosomal RNA
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein
RT-qPCR	real-time quantitative PCR
SAP30	Sin3A-associated protein 30
SAP30L	Sin3A-associated protein 30-like
scaRNA	small Cajal body-specific RNA
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
snoRNP	small nucleolar ribonucleoprotein
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SRP	signal recognition particle
SSC(T)	saline-sodium citrate buffer (with Tween-20 detergent)
SUMO	small ubiquitin-like modifier
TDP-43	TAR DNA-binding protein 43
TER	telomerase RNA
TGF-β	transforming growth factor beta
tRNA	transfer-RNA
UBF	upstream binding factor
UV	ultraviolet
VHL	von Hippel-Lindau tumor suppressor protein
WISH	whole-mount *in situ* hybridization

Tiivistelmä

Tumajyvänen on solun tumassa erottuva rakenne, jossa ribosomien biogeneesi tapahtuu. Ribosomit ovat vastuussa proteiinien tuottamisesta soluissa, ja ne koostuvat RNA:sta (ribosomaalinen RNA, rRNA) ja ribosomaalisista proteiineista. rRNA syntetisoidaan ja sitä muokataan tumajyväsessä, ja myös rRNA:n ja ribosomaalisten proteiinien yhdistäminen kokonaiseksi ribosomiksi alkaa tumajyväsessä. Tämän lisäksi tumajyvänen osallistuu muihin solun toimintoihin, kuten solusyklin säätelyyn, solun kasvuun ja jakautumiseen sekä soluun kohdistuvan stressin tunnistamiseen. Myös erilaisia sairaustiloja on yhdistetty tumajyväseen. Näistä esimerkkejä ovat syöpä ja virusinfektiot. Tumajyvänen ei ole kalvon ympäröimä, joten molekyylit voivat vapaasti liikkua sisään tumajyväseen ja ulos tumajyväsestä. Tumajyvästen luonne on dynaaminen, ja niiden koko ja lukumäärä solua kohden vaihtelevat solun tilan ja solusyklin mukaan. Syöpäsoluilla on usein suurentuneet tumajyväset, mikä on seurausta solun kasvaneesta tarpeesta tuottaa proteiineja, mikä edellyttää ribosomien tehokasta valmistamista tumajyväsessä. Tämä näyttäisi kuitenkin olevan vain osa kokonaisuudesta, sillä myös itse tumajyväselle on ehdotettu roolia syövän kehittymisessä.

Leukemiat eli verisyövät ovat taudinkuvaltaan ja ennusteeltaan kirjava joukko tauteja. Akuuteissa leukemioissa poikkeavia epäkypsiä verisoluja tuotetaan niin runsaasti, että normaali hematopoieesi häiriintyy. Leukemiasolut eivät muodosta kiinteää kasvainta, vaan täyttävät luuytimen ja kulkevat verenkierron mukana. Akuutti lymfoblastinen leukemia (ALL) on lasten yleisin leukemiatyyppi. Nykyisten hoitojen ansiosta hoitotulokset ovat yleisesti kohentuneet, erityisesti pediatrisilla ALL-potilailla, joista lähes 90 % paranee. Tästä huolimatta leukemian syntyyn ja relapsiin vaikuttavat mekanismit ovat vaillinaisesti tunnettuja. Mielenkiintoista on, että historiallisesti niillä ALL:n alatyypeillä, joissa tumajyväset olivat poikkeavia, oli myös huonompi ennuste.

Tämän väitöskirjatyön tavoitteena oli tutkia tumajyväsen roolia akuuteissa lasten leukemioissa. Erityisesti tumajyväsen pieniä RNA-molekyylejä (snoRNA) ja tumajyväsproteiineja tutkittiin eri akuutin leukemian alatyypeissä. Erot snoRNA-

molekyylien ilmentymisessä ja tumajyväsen proteiinikoostumuksessa eri leukemiatyyppien välillä analysoitiin leukemiasolulinjoissa ja potilasnäytteissä. Lisäksi perustettiin menetelmä, jolla tumajyväsproteiinien toimintaa elävässä organismissa on mahdollista tutkia. Tähän käytettiin seeprakalamallia.

snoRNA-molekyylien ilmentyminen analysoitiin solulinjoissa, jotka vastaavat akuutin leukemian päätyyppejä (AML [akuutti myeloinen leukemia], T-ALL ja pre-B-ALL). Tulokset osoittavat, että snoRNA-ilmentyminen vaihtelee eri leukemiatyypeissä, mikä viittaa siihen, että snoRNA-molekyyleillä on leukemiatyyppi-spesifinen ilmentymiskaavio. Useiden yksittäisten snoRNA-molekyylien havaittiin olevan eriävästi ilmentyneitä T-ALL:ssa verrattuna pre-B-ALL:an. Yhdeksän snoRNA-molekyylin ilmentyminen tutkittiin myös kvantitatiivisella PCR:lla, jonka tulokset osoittavat samanlaiset ilmentymiserot. Lisäksi 15 snoRNA-molekyylin ilmentyminen analysoitiin käyttäen laajaa hematologista tietokantaa, johon on kerätty tiedot pediatrisista leukemianäytteistä ja vastaavista terveistä soluista. Useimpien snoRNA-molekyylien ilmentyminen oli tasaista niin leukemianäytteissä kuin terveissäkin, mutta yhden snoRNA:n, SNORD114-3:n, ilmentymistaso oli keskimäärin nelinkertainen viidessätoista näytteessä muihin verrattuna. 13 näistä osoittautui APL-potilaiden näytteiksi. APL eli akuutti promyelosyyttinen leukemia on AML:n alatyyppi. Tämän perusteella SNORD114-3:n korkea ilmentymistaso on merkki APL:sta.

Tumajyväsen proteiinikoostumus analysoitiin leukemiasolulinjoissa, jotka edustivat eri leukemiatyyppejä (AML, T-ALL ja pre-B-ALL). Useita eriävästi ilmentyviä proteiineja tunnistettiin vertailussa eri leukemiatyyppien välillä, ja nämä ilmentymiserot oli mahdollista havaita myös toisella menetelmällä. Laboratoriossamme tunnistetun SAP30L-proteiinin, joka paikantuu sekä tumaan että tumajyväseen, toimintaa tutkittiin seeprakalassa. SAP30L-proteiinin ilmentymisen estäminen seeprakalassa vaikutti alkion kehitystä ohjaaviin signalointireitteihin, mistä seurauksena oli mm. huomattava sydänpussiturvotus, sydämen epämuodostumia, sydämen toimintahäiriöitä sekä punasolujen hemoglobiinitasojen aleneminen.

Näiden tulosten pohjalta voidaan päätellä, että sekä tumajyväsen proteiinien että snoRNA-molekyylien ilmentyminen vaihtelee eri akuuttien leukemiatyyppien välillä. Lisäksi tutkimuksessa käytettiin useita menetelmiä ja välineitä, jotka havaittiin käyttökelpoisiksi tutkittaessa tumajyväsen osasia leukemiassa.

Abstract

The nucleolus, a prominent sub-compartment of the cell nucleus, is the site of ribosome biogenesis. Ribosomes are molecular machines responsible for the production of proteins. The RNA component of the ribosome, ribosomal RNA (rRNA), is synthesized and modified in the nucleolus, where the assembly of rRNA and ribosomal proteins into ribosomes also begins. In addition, the nucleolus is involved in other cellular functions such as regulation of the cell cycle, cell proliferation, and sensing of cellular stress. Also pathological conditions such as cancer and viral infections have been linked to the nucleolus. Since the nucleolus is not membrane-bound, molecules are able to enter and exit it freely. The nucleoli are by nature dynamic, and their number and size vary based on the state of the cell. Cancer cells often display enlarged nucleoli, this resulting from an increased need for protein synthesis, which in turn requires efficient production of ribosomes. This seems, however, to be only part of the truth, as the nucleolus has been held to be implicated in the development of cancer.

Leukemia is a cancer which originates in the blood-forming tissues, and includes a wide spectrum of diseases with different manifestations and outcomes. In acute leukemia, abnormal immature blood cells are produced in large quantities to such an extent that normal hematopoiesis is displaced. The leukemic cells do not form a solid tumor but fill the bone marrow and also circulate in the blood. Acute lymphoblastic leukemia (ALL) is the most common form of the disorder in children. With current treatments outcomes have improved, especially in pediatric ALL, with cure rates close to 90 %. Nevertheless, the mechanisms underlying the development of leukemia and relapse are still insufficiently understood. Historically, ALL subtypes displaying abnormal nucleoli have been associated with poorer prognoses.

The aim of this series was to study the role of the nucleoli in acute pediatric leukemia. More specifically, small nucleolar RNAs (snoRNAs) and nucleolar proteins were studied in different acute leukemia subtypes. The differences in the expression of snoRNAs and in the composition of nucleolar proteins between

leukemia subtypes were analyzed using leukemic cell lines and patient samples. Moreover, a means of investigating the *in vivo* function of nucleolar proteins was set up in the zebrafish.

The expression of snoRNAs was analyzed in cell lines representing the main acute leukemia subtypes (AML (acute myeloid leukemia), T-ALL and pre-B-ALL). The results show differential expression of snoRNAs between the subtypes, which suggests that snoRNAs have a subtype-specific expression pattern. Several individual snoRNAs were found to be differentially expressed between T-ALL and pre-B-ALL. The expression of nine snoRNAs was further analyzed by quantitative PCR, and the results show similar changes in expression. Further, the expression of 15 snoRNAs was analyzed in pediatric leukemia and control samples available in a large hematological dataset. For most snoRNAs, the expression was relatively uniform in both leukemic and normal samples. For SNORD114-3, the expression was consistent among pediatric samples except for 15 cases with expression approximately fourfold higher. When further examined, 13 of these cases turned out to be APL (acute promyelocytic leukemia), a subtype of AML. Thus, overexpression of SNORD114-3 implies APL.

The nucleolar proteomes were analyzed in leukemic cell lines representing AML, T-ALL and pre-B-ALL. Several differentially expressed proteins were identified among the leukemia subtypes, and the results could be reproduced using another method. The function of the nuclear/nucleolar protein SAP30L, which has been identified in our laboratory, was studied in the zebrafish. The depletion of SAP30L affected several signaling pathways that direct early zebrafish development as manifested by marked pericardial edema, deformed cardiac morphology, impaired cardiac function and reduced levels of hemoglobin.

Based on the results, it may be concluded that both nucleolar proteins and snoRNAs are differentially expressed among the acute leukemia subtypes. Furthermore, several tools and methods were applied which proved useful in the study of nucleolar components in leukemia.

1. Introduction

The nucleolus is the sub-compartment of the nucleus where ribosome biogenesis takes place. Ribosomes are responsible for protein production in the cell, as they translate the genetic information of the messenger-RNA (mRNA) into protein sequences. The ribosomes consist of RNA and protein components, and production and modification of the ribosomal RNA components are carried out exclusively in the nucleoli. Moreover, the ribosomes are assembled in the nucleoli with protein components imported from the cytoplasm. In addition to this major function of ribosome biogenesis, a number of other tasks fall to the nucleolus, for example control of cell cycle and proliferation and stress-related events (Boisvert et al. 2007, Sirri et al. 2008). The nucleolus is not bound by a membrane, and molecules can thus freely shuttle between the nucleolus and nucleoplasm. The nucleoli are dynamic, as their size and number vary according to cellular status. Interestingly, it has long been known that cancer cells often display large nucleoli, which has traditionally been attributed to the increased need for protein production in proliferating cancer cells (Montanaro et al. 2008). This appears, however, not to be the whole truth, as the involvement of the nucleoli in the development of cancer has been suggested (Maggi and Weber 2005).

The acute leukemias constitute a type of cancer which target the blood-forming tissues, and lead to production of abnormal blood cells in large numbers. These cells fill the bone marrow and displace normal hematopoiesis, and also circulate in the blood, which leads to the emergence of typical symptoms such as anemia, and a tendency to bleeding and infections. Acute lymphoblastic leukemia (ALL) is the most common type of pediatric leukemia (Inaba et al. 2013). Interestingly, it has long been known that ALL subtypes which display abnormal nucleoli are associated with poorer prognosis. This has also been noted in other types of tumors (Derenzini et al. 2009).

Based on these observations, this present series was undertaken to gain an insight into the potential role of the nucleolus in pediatric acute leukemias. The expression of two different types of molecules present in the nucleolus, small

nucleolar RNAs (snoRNAs) and nucleolar proteins, was studied in various leukemia subtypes. Also, a model system to study the *in vivo* function of nucleolar proteins was set up in the zebrafish.

2. Review of the literature

The central concept in molecular biology is the flow of information from DNA via RNA to protein, these being the active components in the cell. Although this view of gene expression is now known to be incomplete (ENCODE Project Consortium et al. 2012), especially regarding the roles of various RNA molecules, the operations during which the genetic information encoded in DNA is transcribed by RNA polymerase II (RNA pol II) into mRNA, subsequently giving rise to the nascent polypeptide chain via translation by the ribosomes, are still the essential basis of our understanding of cellular function. These operations are carried out by specific molecular machineries, and are under extensive regulation at various levels (Alberts et al. 2002). The nucleolus lies in many respects at the crossroads of these pathways.

2.1 The nucleolus

The nucleolus is the most prominent sub-compartment of the cell nucleus. It was first observed more than 200 years ago, and since then the variable morphology of the nucleoli has been imaged by basic light microscopy. With progress in analytical techniques, more insight into the structure and function of this nuclear body has gradually been gained. The nucleolus is a unique compartment of the cell, as it is not bound by a membrane (Alberts et al. 2002, Sirri et al. 2008).

The major function of the nucleolus is to provide the site where ribosome biogenesis takes place. In fact, in consequence of this function the nucleolus is formed as the various factors assemble together. Each eukaryotic cell has at least one nucleolus, the size and organization of which is linked to ribosome production. The nucleoli are highly dynamic, as their size and number vary according to cell cycle and function. The nucleoli are disassembled at the beginning of mitosis, and reassemble at the end of this process (Hernandez-Verdun 2011). Additional functions assigned to the nucleolus include regulation of mitosis, cell-cycle progression and cell proliferation, acting as a sensor for stress and eliciting stress

responses, and biogenesis of ribonucleoprotein particles. In addition, several lines of evidence associate the nucleolus with human diseases and pathologies (Boisvert et al. 2007).

2.1.1 The structure of the nucleolus

Nucleoli form around the nucleolar organizer regions, NORs, which consist of tandemly repeated clusters of ribosomal DNA (rDNA), giving rise to ribosomal RNA (rRNA). The rDNA repeat consists of a sequence encoding the pre-rRNAs (about 13-14 kilobases, kb), separated by long intergenic spacers (IGSs, approximately 30 kb) which contain all the regulatory elements but are mainly composed of simple sequence repeats and transposable elements (McStay and Grummt 2008). In humans, some 400 copies of the 43 kb rDNA repeat units are distributed along the short arms of the acrocentric chromosomes 13, 14, 15, 21 and 22, in telomere-to-centromere orientation (Boisvert et al. 2007, Sirri et al. 2008). Thus, in a diploid genome altogether 10 NORs are present, although in many cell types only some of them are transcriptionally active. The active and silent NORs display distinct epigenetic states (McStay and Grummt 2008). In addition to NORs, the nucleoli contain a complex set of proteins, hundreds of snoRNAs and rRNAs.

The nucleolus consists of three structurally and functionally distinct compartments: the fibrillar center (FC), the dense fibrillar component (DFC), and the granular component (GC) (see Fig.1). The FCs are entirely or partially surrounded by DFC, and the GC composed of granules (pre-ribosomal particles of variable maturation state) envelops the FC and DFC. The transcription of rDNA takes place either in the FCs or in the border between FC and DFC, and the FCs abound with components of the rDNA transcription machinery. The nascent transcripts emerge at the border of FCs and DFC, and reside in the DFC, where the 47S pre-transcript processing begins. During this processing the pre-rRNAs gradually migrate towards the GC, where the intermediate and late processing stages and ribosome subunit assembly occur. The proteins involved in the specific stages of rRNA processing reside in the corresponding nucleolar areas. The cellular status of metabolism and growth affects nucleolar activity, which is in turn reflected in the size and morphology of the nucleolus (Bartova et al. 2010, Boisvert et al. 2007, Boulon et al. 2010, Sirri et al. 2008). In conclusion it can be stated that the organization of the nucleolus results from ribosome biogenesis.

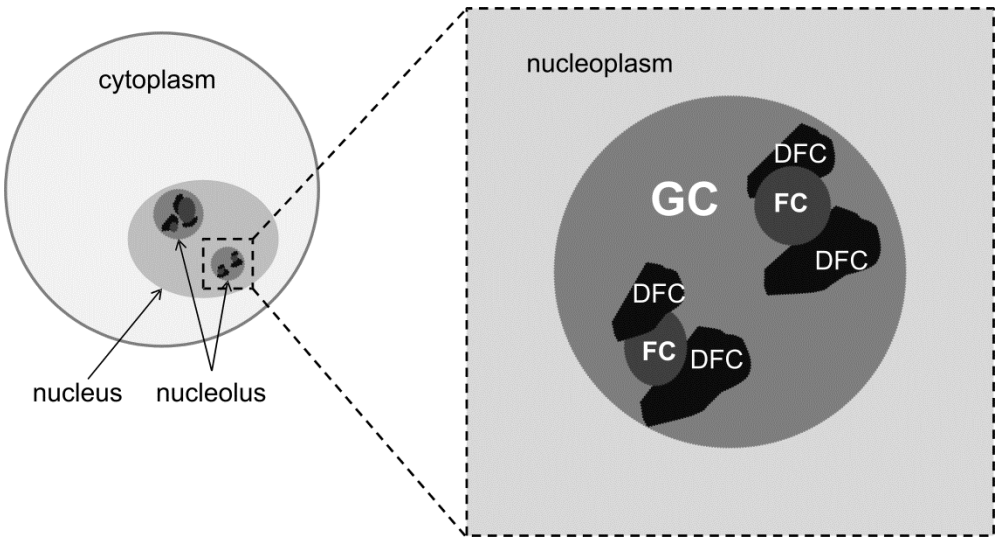

Figure 1. The cell and the nucleolus. A schematic illustration of a cell with two nucleoli present in the nucleus. One of the nucleoli is shown magnified on the right, with two fibrillar centers (FC) surrounded by a dense fibrillar component (DFC) embedded in the granular component (GC).

2.1.2 The nucleolar proteome

The dynamic nature of the nucleolus is reflected in its protein content, which has proved to be unexpectedly complex (Andersen et al. 2002). In order to gain an overall picture of the functions the nucleolus is involved in, the proteins identified in the nucleolar proteome were divided into nine categories: ribosomal proteins, ribosome biogenesis, chromatin structure, mRNA metabolism, translation, chaperones, fibrous proteins, others, and unknown function (Coute et al. 2006). In the nucleolar proteome reported by Andersen et al. (Andersen et al. 2005), ribosomal and ribosome biogenesis-associated proteins make up about one third of the total number of proteins, considerably less than reported in earlier analyses (Andersen et al. 2002, Scherl et al. 2002). Also, the relative proportion of proteins categorized into mRNA metabolism or "other" was higher in the 2005 study, this probably reflecting the greater total number of proteins identified in the analysis (Andersen et al. 2005, Coute et al. 2006). As these observations would imply, the nucleolus is involved in many cellular functions. More recently, an extensive

proteomic analysis aiming to uncover the nucleolar proteome revealed that over 4 500 proteins are localized in the nucleolus (see http://www.lamondlab.com/NOPdb3.0/ for updated list) (Ahmad et al. 2009).

The dynamic nature of the nucleolar proteome is also manifested under challenging cellular conditions. When HeLa cells were stressed with actinomycin D, which inhibits transcription by all RNA polymerases, some interesting observations were made (Andersen et al. 2005). First, the changes in nucleolar protein levels were due mainly to redistribution, while the overall levels of protein remained unchanged. Second, factors involved in RNA processing, rDNA transcription, and ribosomal proteins exited the nucleolus following an rDNA transcription halt, suggesting that their presence in the nucleolus is closely related to ribosome synthesis. Third, a subset of proteins accumulated in the nucleolus, some up to tenfold, indicating that not all redistribution of nucleolar proteins is dependent upon ribosome synthesis. The nucleolus is thus not merely a ribosome biogenesis machine which breaks down in the absence of rDNA transcription. Furthermore, since the nucleolar proteome varies in response to different stimuli, it would appear that there is no one complete proteome, but instead several overlapping sets of proteins relevant for particular conditions in the cell. It is important to observe that some proteins are also present in other cellular locations, and some only accumulate in the nucleolus transiently (Andersen et al. 2005).

The above notwithstanding, as a consequence of the major activity of the nucleolus, ribosome synthesis, many proteins residing in the nucleolus are involved in various steps of this process. Proteins associated with rDNA transcription include the RNA polymerase I itself, a 14 subunit enzyme, a number of transcription factors, the upstream binding factor (UBF), the selectivity factor SL1, consisting of at least four subunits, including the TATA-binding protein (TBP), and TIF-IA (RRN3 in yeast), which interacts directly with the RNA polymerase (Schneider 2012). Several exo- and endonucleases are necessary for pre-rRNA cleavages, and the rRNA-modifying proteins are also present in ample numbers. In order to identify the proteins necessary for rRNA processing, over 600 nucleolar proteins were screened with small interfering RNAs (siRNAs), and a total of 286 proteins were found to be required, the majority of them apparently taking part in the early processing steps (Tafforeau et al. 2013). The small nucleolar ribonucleoprotein (snoRNP) complexes carry out the covalent modifications of rRNA, methylation and pseudouridylation. The protein components of these

complexes include fibrillarin, NOP56, NOP58, 15.5 kDa, dyskerin, GAR1, NHP2, and NOP10 (Reichow et al. 2007). The snoRNAs and their functions are described in greater detail in section 2.1.3 and Fig.2.

The ribosomal proteins (RPs) are transported to the nucleolus, where they are assembled with the rRNAs to produce the large (L) and small (S) ribosomal subunits. The proteins are grouped into RPL or RPS, depending on the subunit they reside in. The precise proportions of the altogether 78 RPs incorporated in the subunits is important, as deletions or overexpression of these proteins affect ribosome production as a whole. Additional functions of the RPs include roles as rRNA chaperones and stabilizing factors. Furthermore, extraribosomal functions such as monitoring correct ribosome assembly, or in growth regulation, have been reported (Lafontaine and Tollervey 2001, Ruggero and Pandolfi 2003, Warner and McIntosh 2009).

The multifunctional role of nucleolar proteins is exemplified by two major nucleolar proteins. Nucleophosmin (NPM, B23), encoded by the *NPM1* gene, is a histone chaperone protein involved in nucleosome assembly and disassembly, but it also exerts its chaperone activity on a variety of targets. NPM is ubiquitously expressed, and most of the protein resides in the nucleolus, although a small proportion of its molecules shuttle constantly between the cytoplasm and the nucleus. NPM is also involved in rDNA transcription and rRNA processing. There are numerous reported interaction partners for NPM, including both RNA and protein molecules. Through this complex interaction network NPM contributes to the promotion of cell growth and survival. It is thus not surprising that NPM is often implicated in cancer, especially in hematopoietic malignancies (Colombo et al. 2011, Grisendi et al. 2006). Nucleolin (NCL, C23) is an abundant nucleolar protein which has effects on rDNA transcription, rRNA processing, and ribosome assembly. Like NPM, NCL also evinces histone chaperone activity, and has been linked to chromatin remodeling. It is also present in other cellular locations, and binds both RNA and DNA. In addition, NCL affects the mRNA stability and translation of several proteins such as p53, and plays a role in viral infection, cancer and inflammation (Abdelmohsen and Gorospe 2012, Mongelard and Bouvet 2007). Several other major multifunctional proteins of the nucleolus have also been identified.

It has been shown that most nucleolar proteins, including those involved in ribosome biogenesis, shuttle between the nucleolus and the nucleoplasm in a continuous cycle (Sirri et al. 2008). The nucleolar localization signal (NoLS) which causes a protein to be targeted to the nucleolus is not well defined, as no clear consensus sequence has been found among several NoLSs described. Furthermore, proteins found to be located in the nucleolus do not share a common targeting motif similar to the nuclear localization signal (NLS) (Andersen et al. 2002). Nevertheless, when the reported NoLSs were analyzed, the following observations were made: 1) about half of the residues in the NoLS are basic (arginine or lysine), 2) its secondary structure is an α helix or a coil found predominantly on the protein surface, and 3) its location is close to one or other terminus of the protein (Scott et al. 2010).

Since the nucleolus is not membrane-bound, it seems likely that interactions between protein molecules regulate their nucleolar localization. One theory places NPM in the role of a nucleolar transporter, as it constantly travels from the nucleolus to the cytoplasm and back, and has been shown to interact with numerous other nucleolar proteins (Sirri et al. 2008). Another cellular regulation strategy is proposed for the nucleolar detention pathway (NoDP), which leads to the immobilization of proteins within the nucleolus (Audas et al. 2012b). NoDP uses the interaction of the nucleolar detention signal (NoDS) of the target proteins with members of a novel class of inducible long non-protein-coding RNA (lncRNA) transcripts, which originate from the IGS repeats of rDNA. This interaction captures target proteins such as VHL, DNA methyltransferase DNMT1, and the delta subunit of DNA polymerase (POLD1), in the nucleolus, thereby preventing them from executing their normal function (Audas et al. 2012a).

2.1.3 Small nucleolar RNAs (snoRNAs)

The rRNAs are the most abundant RNAs in the cell, accounting for approximately 80 % of total RNA in an active cell. The two main types of modification to rRNA are 2´-O-ribose methylation and pseudouridylation. Each modification is carried out by a specific snoRNP complex consisting of proteins and a snoRNA. The snoRNAs are an evolutionarily ancient group of non-protein-coding RNAs (ncRNAs) (Boisvert et al. 2007, Kiss 2002). They subsist in two classes, the box C/D and box H/ACA snoRNAs. The nomenclature derives from the distinctive,

evolutionarily conserved sequence elements present in the snoRNAs, which coordinate interaction with the protein components of the snoRNP complex. snoRNAs of each type fold into a specific secondary structure. As the name snoRNA implies, these RNA molecules are small (most of them are 60 - 300 nucleotides in length), and reside in the nucleolus (Kiss 2002). Currently there are 382 identified snoRNAs in the snoRNA-LMBE-database (retrieved 20.9.2013 (Lestrade and Weber 2006)), 274 of them being box C/D type and 108 box H/ACA type, with an additional 25 small Cajal body-specific RNAs (scaRNAs), which also contain box H/ACA and/or box C/D sequence elements (Darzacq et al. 2002). The classical function of the snoRNAs is to target the appropriate nucleotide for modification by complementary base pairing with the target RNA, thus positioning the snoRNP complex correctly. snoRNAs also guide modification of other RNAs such as transfer-RNAs (tRNAs), and small nuclear RNAs (snRNAs) involved in mRNA splicing. A subset of snoRNPs from each class is involved in pre-rRNA cleavage events (Kiss 2002, Reichow et al. 2007). Interestingly, additional functions have also been proposed. This expands the straightforward conception of snoRNAs as being involved only in housekeeping tasks to broader roles in cell fate, oncogenesis, and cellular behavior (Williams and Farzaneh 2012).

Most human snoRNAs are encoded in the introns of protein-coding genes, the host genes, following the "one snoRNA per intron" rule, and are transcribed by RNA pol II. As the host pre-mRNA is spliced, the snoRNA-containing introns are excised and undergo further processing. Correct processing and stabilization of snoRNAs require the ordered recruitment of the four box C/D and four box H/ACA protein components. Box C/D snoRNA genes usually reside in relatively short introns, about 80 nucleotides upstream of the 3' splice site, whereas box H/ACA snoRNA genes are often located in introns which are longer than average (Kiss et al. 2006, Richard and Kiss 2006). Some snoRNA host genes appear to be non-protein coding, i.e. the end product is an RNA transcript, the introns of which contain the snoRNAs. Such genes usually encode several snoRNAs (Dieci et al. 2009). Recently, snoRNA-independent effects have been described for the host ncRNAs (Askarian-Amiri et al. 2011, Mourtada-Maarabouni et al. 2009). More than 90 % of human snoRNAs reside within introns of protein-coding or non-coding genes. In addition, there is an over-representation of genes involved in ribosome biogenesis and translation among the protein-coding host genes, suggesting a coordinated regulation of expression of the components involved in these

processes. However, more diverse gene locations and expression strategies have been reported for other snoRNAs. For instance, some snoRNAs have an intergenic location and are transcribed from independent promoters (Dieci et al. 2009). Interestingly, it has been reported that snoRNAs may also be epigenetically regulated in cancer, as CpG island hypermethylation leads to inactivation of the associated snoRNAs (Ferreira et al. 2012). Apart from these observations, relatively little is known as to the regulation of snoRNA expression.

The box H/ACA snoRNAs are a large family of ncRNAs present in all eukaryotes and archaea, and they associate with four evolutionarily conserved proteins to create the box H/ACA snoRNP complex. The protein components include dyskerin (DKC1), which maintains the catalytic activity, NHP2, NOP10 and GAR1 (see Fig. 2A). The box H/ACA snoRNP conveys site-specific pseudouridylation on RNA, that is, the isomerization of uridine residues to the C5-glycoside isomer pseudouridine. Pseudouridine, which is considered to be the most common post-transcriptionally synthesized modified ribonucleotide in cellular RNAs, contains an additional hydrogen bond donor compared to uridine (Kiss et al. 2010). Box H/ACA snoRNAs operate mainly in the nucleolus, where they target rRNA, but a specific group, scaRNAs, which target the spliceosomal snRNAs, operate in the nucleoplasmic Cajal bodies. The box H/ACA snoRNAs contain two sequence elements, the box H (AnAnnA) and box ACA, which remain single-stranded when the RNA folds into the typical secondary structure of two hairpins (see Fig. 2A). In addition to pseudouridylation, box H/ACA snoRNAs can serve as miRNA-precursors, and contribute to telomerase activity (Kiss et al. 2010). Telomerase is an RNP with reverse transcriptase activity, and maintains telomere length at chromosome ends. Telomerase contains an RNA component, TER, which acts as template for the telomeric repeat sequence DNA, and is structurally also a box H/ACA snoRNP containing dyskerin (Collins 2006). In conclusion, the box H/ACA snoRNPs are involved in three important cellular processes: protein synthesis, mRNA splicing, and maintenance of genomic integrity.

The box C/D snoRNAs guide 2´-O-ribose methylation of nucleotides in RNA. They associate with four protein partners to form the functional box C/D snoRNP: fibrillarin, which holds the methyltransferase activity, NOP56, NOP58, and 15.5kDa (see Fig. 2B) (Kiss 2002). The sequence element motifs of box C/D snoRNAs are RUGAUGA (box C; R is a purine) near the 5' end of the snoRNA,

and CUGA (box D) close to the 3' end. When the box C/D snoRNA acquires the classical stem-internal loop-stem structure, box C and D motifs are brought together by base pairing of the 5' and 3' ends of the snoRNA (see Fig. 2B). Many box C/D snoRNAs contain a second set of the C and D box sequence elements in the central region of the snoRNA named C' and D'; these are usually less conserved (see Fig. 2B) (Kiss 2002, Reichow et al. 2007). According to the current understanding, eukaryotic box C/D snoRNPs are assumed to adopt a pseudo-symmetric architecture where the 15.5kDa protein associates with the C/D site of the snoRNA and initiates snoRNP assembly. NOP56 and NOP58 recognize the box C and C' elements and one copy of fibrillarin is associated with each D and D' box, as required for catalytic activity (see Fig. 2B) (Reichow et al. 2007).

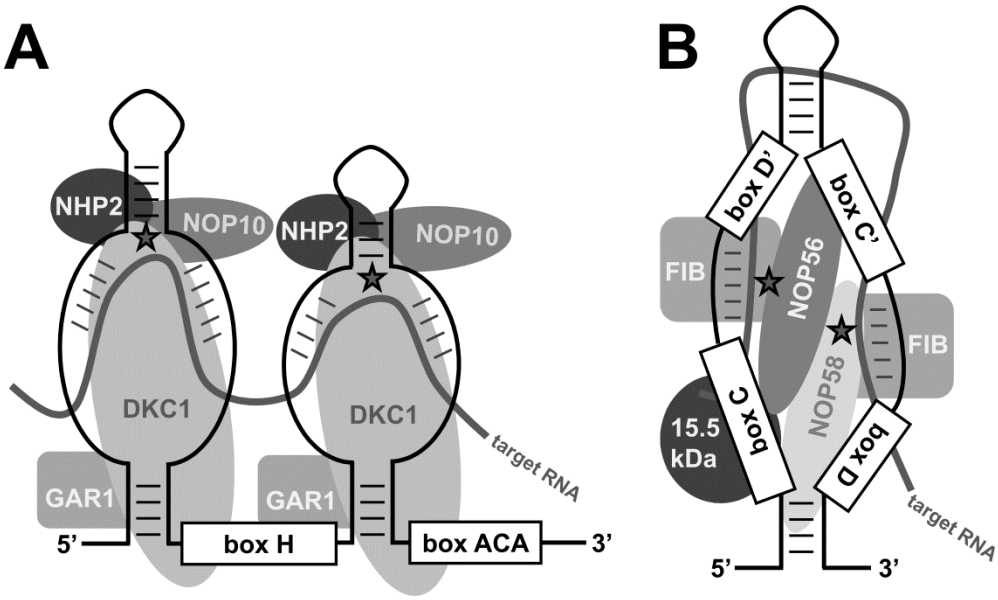

Figure 2. A schematic illustration of snoRNPs with target RNAs. snoRNAs are in black with box motifs and base pairing shown. Star indicates the site of modification. A. The box H/ACA snoRNP. The guide sequences for pseudouridylation are located in the so-called pseudouridylation pockets of both hairpin motifs, and thus one box H/ACA RNA can guide two modifications, carried out by separate sets of box H/ACA proteins. B. The box C/D snoRNP. The modification sites for 2'-O-methylation are located exactly 5 nucleotides upstream of the box D or D' sequence, and are determined by 10-20 nucleotide base pairing between the box C/D snoRNA and the target RNA. Modified from Reichow et al. 2007.

In human rRNA, there are approximately 90 pseudouridylated residues and over 100 2'-O-methylated residues, and in both ribosomal subunits the modifications occur at evolutionarily conserved sites correlating with functionality. The chemical properties of the modified nucleotides are altered, affecting e.g. steric properties, hydrogen-bonding potential, structural rigidity and base stacking. In each case, the structural context determines the overall effects on structural and thermodynamic features (Decatur and Fournier 2002). The question of the significance of the rRNA modifications is one of the oldest in RNA science. They have been considered beneficial, as the early studies conducted in e.g. yeast have revealed: blocking the methylation and pseudouridylation activities exert a strong negative effect on growth rate. In addition, the modifications are located at sites of known or predicted functional importance (Decatur and Fournier 2003). Several pieces of evidence have since indicated the correct rRNA modifications to be significant to both rRNA synthesis and ribosome function. For example, mis-targeted methylation in rRNA results in marked impairment in ribosome activity and reduced translation rates (Liu et al. 2008), the loss of modifications at critical rRNA sites impairs translation and causes delays in pre-rRNA processing (Liang et al. 2009), and defects in rRNA pseudouridylation affect ligand binding and translational fidelity of the ribosome (Jack et al. 2011).

An interesting snoRNA subgroup are the so-called orphan snoRNAs, which lack a known RNA target. One suggestion regarding their function would implicate them in alternative splicing of mRNAs (Bazeley et al. 2008). Indeed, recent observations on small RNAs derived from box C/D snoRNAs link these molecules to alternative splicing events (Scott et al. 2012). Another proposed novel function for snoRNAs is to give rise to shorter RNAs such as microRNAs (miRNAs), via nucleolytic processing. miRNAs regulate the translation of specific mRNAs by binding the 3'UTR region. This has thus far been documented for snoRNA ACA45 (Ender et al. 2008), some box H/ACA snoRNAs (Scott et al. 2009), and some box C/D snoRNAs (Brameier et al. 2011, Ono et al. 2011). Some of these molecules appear to function as a snoRNA, and, after processing, as a miRNA (Scott and Ono 2011). Given the accumulating data on unorthodox snoRNA functions, it is also possible that only a fraction of the snoRNAs, the abundant snoRNAs most likely involved in the classical rRNA modification task, have so far been identified. With more sensitive techniques more novel snoRNAs with different functions are likely to be discovered, expanding our general view of snoRNAs and their functions.

2.1.4 Nucleolar functions

2.1.4.1 Ribosome biogenesis

Ribosomes are the protein factories of the cell responsible for interpreting the nucleotide sequence of the mRNA and converting it in a codon-wise manner into a corresponding amino acid chain linked with peptide bonds. The mature ribosome is composed of two ribonucleoprotein subunits, the large and the small. The small subunit, 40S in eukaryotes, has the mRNA decoding function, whereas the large subunit, 60S, exercises the catalytic function needed for the formation of peptide bonds between amino acids. Both subunits consist of rRNA and RPs: the large 60S subunit is made up of 5S, 5.8S and 28S rRNA and 46 RPs, and the small 40S subunit is composed of 18S rRNA and 32 RPs (Lafontaine and Tollervey 2001).

The first step in attributing the production of ribosomes to the nucleolus was the simple observation that the ribosomal genes are located in the nucleoli. The size and organization of the nucleoli was observed to be linked to the activity of ribosome biogenesis, as nucleolar prominence was seen in active cycling cells, whereas the nucleoli were of limited size in terminally differentiated cells such as lymphocytes (Sirri et al. 2008). During the first step of ribosome biogenesis, the rDNA is transcribed by RNA polymerase I (RNA pol I) into a 47S rRNA precursor transcript, which is next processed in a series of cleavages aided by various endo- and exonucleases, and modified by snoRNP complexes to produce the mature 28S, 18S and 5.8S rRNAs. Over 200 factors participate in this processing of rRNA. The fourth rRNA, the 5S rRNA, is transcribed by RNA polymerase III (RNA pol III) from a separate gene cluster, and does not undergo chemical modification. Even though the ribosome subunits contain both RNA and protein, the sites evincing catalytic activity comprise mostly rRNA (Alberts et al. 2002). The rRNAs then associate with the protein components to produce the ribosomal subunits and are transported to the cytoplasm. The 40S and 60S ribosome subunits are exported from the nucleus via separate routes. The process is better understood for pre-60S and involves recruitment of various export receptors and adaptors in the nucleoplasm before translocation through the nuclear pore complex channel to the cytoplasm after which the auxiliary components are removed (Kohler and Hurt 2007). The functional, translationally active ribosome is

formed in the cytoplasm. rRNA synthesis is the rate-limiting step in ribosome biogenesis (Montanaro et al. 2012).

It is worth noting that several networks must co-operate to produce the ribosome subunit. rDNA transcription, rRNA processing, and ribosomal subunit assembly are separate processes with distinct machineries operating in sequential order in a temporally and spatially defined manner. To some extent these processes interact, influence each other, and share regulatory elements, but are, all in all, under complex regulation. For example, the transcription of rDNA is under epigenetic regulation, as the acetylation, methylation and ubiquitination status of the histones determines the accessibility of the transcription machinery to the chromatin (McStay and Grummt 2008). Additionally, rRNA processing covers several functions including a series of pre-rRNA cleavages as well as covalent modification of the maturing rRNA molecules (Boisvert et al. 2007, Hernandez-Verdun 2011, Sirri et al. 2008).

2.1.4.2 RNA modification

As described above, the rRNAs undergo covalent modification in the nucleolus during maturation, and such modification is essential for the correct function of the ribosome. In addition to rRNA modification, the nucleolus is involved in the maturation and processing of other cellular RNA types. For example, it participates in the modification and assembly of the spliceosomal small nuclear RNPs (snRNPs), telomerase, several other small RNA species, the signal recognition particle (SRP), and miRNAs. snRNAs such as U2, U4, U5 and U6 are modified by snoRNPs, similarly to rRNA, by inducing methylation and pseudouridylation. The SRP complex is responsible for the detection of the N-terminal signal and correct targeting of the nascent peptides to the endoplasmic reticulum. It contains 6 proteins and a 300 nucleotide-long RNA, which transit through the nucleolus before transportation to the cytoplasm (Boisvert et al. 2007, Gerbi et al. 2003).

2.1.4.3 Cell cycle and the nucleolus

The cell cycle affects the nucleolus and vice versa. The most obvious example is the cellular need for functional ribosomes, with the greatest demand for rDNA transcription in the S and G2 phases, whereas during mitosis, when ribosome

production ceases, the nucleolus is dispersed. On the other hand, some regulators of the cell cycle reside in the nucleolus and interact with the nucleolar proteins.

In higher eukaryotes, the progression of the cell cycle and ribosome biogenesis are closely linked. The production of ribosomes begins at the end of mitosis, is increased during G1-phase, peaks in G2, and ceases in the prophase of mitosis. The RNA pol I machinery remains associated with the rDNA of the active NORs, whereas factors involved in the processing of the rRNA are redistributed from NORs. At the end of the prophase, when the chromosomes are condensed and the nuclear envelope disintegrates, the nucleolus is no longer visible. The perichromosomal compartment, where the condensed chromosomes reside, also holds some protein complexes identified in the DFC and GC, suggesting that partial building blocks for the future nucleoli are stored during mitosis. The shutdown of rDNA transcription is mediated by phosphorylation of the involved proteins, and similarly, the phosphorylation status of the nucleolar processing proteins is modified during the prophase, leading to e.g. weakened capacity for interaction and release from the nucleolus. Generally, the disassembly of the nucleoli is faster than its reassembly. At the end of mitosis, the rDNA transcription starts at NORs, and rRNA processing machineries become translocated in foci called prenucleolar bodies (PNBs). From PNBs these components move in sequential order to active rDNA transcription sites, and the compartmentalized nucleoli with FC, DFC and GC are formed (Boisvert et al. 2007, Hernandez-Verdun 2011).

It is obvious that the nucleoli both receive signals and respond to them. Although the nucleoli are reassembled at the end of mitosis and the beginning of the interphase, they remain dynamic throughout the latter. Various cycling proteins associate with the nucleoli in a cell cycle-specific manner, i.e. during certain stages of the cell cycle. Some of these proteins are involved in post-translational modification of proteins, e.g. sumoylation and phosphorylation, which are dynamic and reversible processes, and can influence numerous activities of the cell (Boisvert et al. 2007). For example, SENP5, a SUMO-specific protease responsible for desumoylation, has a predominantly nucleolar localization. When SENP5 is knocked down by RNAi, the cells exhibit aberrant nuclear morphology and defects in cell division (Di Bacco et al. 2006). An example of nucleolar localization of components participating in reversible protein phosphorylation is provided by the protein phosphatase 1 (PP1), a ubiquitous serine-threonine phosphatase which

regulates several cellular functions. PP1 has three active isoforms with distinct localization patterns, and during the interphase a pool of the PP1γ isoform accumulates in the nucleoli. When the cell enters mitosis, the PP1γ disperse in the cytoplasm and also associate with the kinetochores (Trinkle-Mulcahy et al. 2007). From the cytoplasm, PP1γ relocalize to the chromosomes and participate in chromatin condensation and remain associated with chromatin during the following interphase, and accumulate again in the nucleoli (Vagnarelli et al. 2006). In this regard, the nucleolus takes part in the segregation of chromosomes and cytokinesis (Trinkle-Mulcahy and Lamond 2006). Sequestration to the nucleolus is also a mechanism used to control other cellular activities. For example, the telomerase reverse transcriptase, an RNP which extends the telomere repeats at the ends of the chromosomes, remains in the nucleoli until telomere replication in the late S phase. Of note, this cell-cycle-dependent nucleolar localization of telomerase is not observed in transformed cells (Wong et al. 2002).

2.1.4.4 Cellular stress and the nucleolus

Typically, stress-induced signaling in the cell results in cell cycle arrest or apoptosis. The cell responds to stress rapidly, and seeks to compensate by adjusting its metabolism. The outcome depends on the cell's ability to recover and the severity of the insult. In addition, the major nuclear operations, transcription and replication, as well as energy-consuming ribosome biogenesis, are often inhibited, which is reflected in the organization of the nucleus (Alberts et al. 2002). The stress the cell experiences can come from a variety of sources. These include factors or chemical compounds inducing DNA damage, changes in temperature, hypoxia, nutrient stress, and viral infection, all of which have been shown to cause nucleolar disruption (Boulon et al. 2010). Interestingly, two agents inducing DNA damage, ultraviolet (UV) and ionizing radiation (IR), lead to different nucleolar responses. UV induces structural reorganization in the nucleolus, whereas IR does not. The nucleolar proteomic changes induced by IR are also more subtle and temporally restricted as compared to those caused by UV. Such observations pinpoint the nucleolus as a sensitive stress sensor able to respond in a damage-specific manner (Moore et al. 2011). The nucleolus appears to be a coordination center for stress response, as the organization and the protein content of the nucleolus is altered under stress stimuli, the signaling pathways exploiting nucleolar accumulation and/or release of proteins in response to these (Boulon et al. 2010).

p53 is a key player in the nucleolar response to stress. Since p53 is considered to be the "guardian of the genome", multiple independent and co-dependent pathways control its activities. Here, the nucleolar perspective on p53 is chosen. In normal conditions, the levels of p53 are kept low, but under stress, it becomes stabilized and active, allowing the induction of the expression of target genes such as p21 (contributing to cell cycle arrest), and Bax and Puma (driving apoptosis) (Lee and Gu 2010, Toledo and Wahl 2006). Interestingly, nucleolar integrity is a prerequisite for p53 to be maintained at low levels in the cell (Boulon et al. 2010, Burger and Eick 2013). The main regulator of p53 is the E3 ubiquitin ligase Hdm2. Under normal conditions, p53 interacts with Hdm2, which leads to its ubiquitination, subsequent export from the nucleus and degradation (Toledo and Wahl 2006). This process requires p53 transit through an intact, functional nucleolus (Boyd et al. 2011). Various stress signals increase the expression of the predominantly nucleolar tumor-suppressor protein p14ARF, which then sequesters Hdm2 and prevents its interaction with p53, this contributing to p53 stabilization (Sherr and Weber 2000). Furthermore, the stability and localization of p14ARF are regulated by NPM in a p53-independent manner (Sherr 2006). RPs also play a role in p53 regulation. Following stress, ribosome biogenesis is disrupted and ribosomal proteins are released from the nucleolus. The RPs L5, L11, L23 and S7 have been shown to interact with Hdm2 in the nucleoplasm, leading to p53 stabilization (Boulon et al. 2010, Zhang and Lu 2009). For RPL11, this function is intensified by stress (Sundqvist et al. 2009). Furthermore, the RPL11-Hdm2 control of p53 activation is regulated by the nucleolar protein PICT1, mainly via binding to RPL11 and retention of it in the nucleolus (Sasaki et al. 2011). The major nucleolar protein NPM has also been shown to protect p53 from Hdm2-mediated degradation by interacting with Hdm2 (Kurki et al. 2004). p53 is also able to disrupt the SL1-UBF interaction at rDNA transcription initiation, causing RNA pol I inhibition and a subsequent decrease in ribosome biogenesis (Zhai and Comai 2000). Taken together, complex interaction networks involving p53, nucleolus and various (nucleolar) proteins operate in the cell under normal circumstances and during stress.

2.1.5 The nucleolus and diseases

As described above, many of the nucleolar functions are essential for cell growth and survival. Accordingly, disturbance of these functions can be manifested at the level of the whole organism. In a recent analysis, factors necessary for ribosome assembly were identified. Interestingly, 38 % of the genes in question have also been implicated in diseases, particularly in cancer (Tafforeau et al. 2013). An early indication of snoRNAs having a role in human disease was provided by the Prader-Willi syndrome (PWS), an interesting multisystem disorder. The condition is characterized by various clinical symptoms such as hypotonia, delayed development, behavioral problems, hypogonadism, scoliosis, and hyperphagia, which can lead to morbid obesity. PWS is caused by the absence of paternally expressed imprinted genes in the chromosomal region 15q11.2-q13. There are several imprinted genes in this region, but the majority of the clinical features can be narrowed down to the SNORD116 gene cluster of 29 snoRNAs. All these snoRNAs are orphans, and have been proposed to take part in splicing events (Cassidy et al. 2012, Williams and Farzaneh 2012).

The human conditions most commonly associated with the nucleolus are cancer and ribosomopathies. In addition, the nucleolus is reported to play a role in viral infections, as discussed below.

2.1.5.1 Cancer and the nucleolus

Long before the attribution of ribosome biogenesis to the nucleolus, the link between unusually large nucleoli and the malignant nature of cells was observed. The early observations concerning the number of nucleoli, the activity status of the cell, and the size of the cell are still valid, and can be utilized as nucleolar indicators of proliferation and tumorigenesis (Maggi and Weber 2005). Although nucleolar hypertrophy and functional up-regulation are considered to be common characteristics of cancer cells, this does not universally apply, as some cancer cells display lower nucleolar size and activity compared to corresponding normal cells (Montanaro et al. 2008). More precisely, the activity of rRNA transcription and nucleolar size are inversely related to the doubling time in cancer cells (Derenzini et al. 1998), and nucleolar size can reliably indicate the rapidity of cancer cell proliferation (Derenzini et al. 2000). Clinically, the presence of enlarged nucleoli,

"prominent nucleoli", is one of the nuclear parameters used to grade a tumor. As such, nucleolar size is not a diagnostically applicable feature. Nevertheless, since nucleolar size is affected by several important characteristics of the cancer cell, for example the proliferation rate and the mutation/inhibition status of p53 and Rb, it is well suited for prognostic purposes. Indeed, the nucleolar parameter has proved to be a relevant prognostic factor for many types of tumors: the larger the nucleoli, the worse the prognosis (Derenzini et al. 2009). This is also the case in acute lymphoblastic leukemia (ALL). The French-American-British (FAB) classification system based on cytomorphology and cytochemistry has traditionally been used to categorize leukemias. Interestingly, the types L2 and L3, which display abnormal nucleoli, are associated with poorer outcomes (see Table 1).

Table 1. The FAB classification of acute lymphoblastic leukemia.

Cytologic feature	***L1***	***L2***	***L3***
Cell size	Predominantly small	Large heterogeneous	Large homogeneous
Nuclear chromatin	Homogeneous	Variable	Stippled homogeneous
Nuclear shape	Not visible or small conspicuous	Commonly irregular clefting and indentation	Regular (oval to round)
Nucleoli	Regular; occasional clefting	One or more, often large	One or more, prominent
Amount of cytoplasm	Scanty	Variable, often moderately abundant	Moderately abundant
Basophilia of cytoplasm	Slight or moderate	Variable, deep in some	Very deep
Cytoplasmic vacuolation	Variable	Variable	Often prominent

Modified from Pui 2006.

Traditionally, the alterations observed in the nucleoli of highly proliferative tumor cells have been attributed to an increased requirement for protein synthesis, and therefore increased ribosome biogenesis. On the other hand, accumulating evidence suggests the involvement of qualitative and quantitative changes in ribosomes, leading to reduced tumor suppressor potential in the cell and subsequently to an increased risk of cancer onset (Montanaro et al. 2012). For

example, an increase in the ribosome biogenesis rate contributes to more aggressive phenotype of breast cancer cells (Belin et al. 2009). Indeed, changes in ribosome biogenesis seen in several human pathological conditions have been found to be associated with an increased cancer risk (see section 2.1.5.2 below).

As described above, under normal conditions the nucleolar functions are controlled by several proteins which regulate cell cycle and proliferation. In tumors, the same proteins have a pro-oncogenic function which is further activated, or a tumor-suppressive function which is inactivated, and these changes lead to the loss of normal control over proliferation and up-regulation of ribosome biogenesis, both essential for tumor progression (Montanaro et al. 2012). The major tumor suppressor proteins are p53, PTEN (phosphatase and tensin homolog), and Rb (retinoblastoma protein), which down-regulate ribosome biogenesis in the nucleolus under normal circumstances, and the activities of which are frequently lost in tumor cells. p53, Rb and PTEN are all able to repress rDNA transcription by RNA pol I, Rb by binding to UBF, and p53 and PTEN by disrupting the SL1 transcription initiation complex. In addition, p53 and Rb inhibit 5S rRNA transcription by RNA pol III (Montanaro et al. 2012, Ruggero and Pandolfi 2003). Interestingly, p53 has been shown to directly repress the expression of fibrillarin, the box C/D snoRNP methyltransferase. In the absence of p53, the expression of fibrillarin is increased, this leading to modification in the methylation pattern of rRNA. Ribosomes with altered rRNA methylation have lower translational fidelity and are more prone to initiate the translation of key oncogenic mRNAs from internal ribosomal entry sites (Marcel et al. 2013). When the regulatory roles of p53 described elsewhere and here are taken together, the nucleolus would appear to be essential for p53 function as the guardian of the genome, and thus highlights the role of the nucleolus in tumorigenesis (Vlatkovic et al. 2013).

The prominent proto-oncogene c-Myc regulates ribosome biogenesis at various levels: 1) increasing RNA pol I activity by recruiting the SL1 complex to the promoter, 2) increasing RP production by stimulating RNA pol II transcription, and 3) enhancing RNA pol III activity (Montanaro et al. 2012). Of note, NPM plays an essential role in localizing c-Myc in the nucleolus and in c-Myc-mediated induction in rDNA transcription (Li and Hann 2013). NPM is further implicated in cancer in that it is often overexpressed. NPM is involved in various cellular processes and has several interaction partners. This contributes to the complex role of NPM in tumorigenesis, which can be either oncogenic or tumor-suppressive,

depending on cellular type and context. Especially in hematological malignancies, NPM is often involved in chromosomal translocations or mutations. In AML, a cytoplasmic mutant of NPM, NPMc+, is a common finding and is thought to contribute to disease development in several ways (Colombo et al. 2011, Hein et al. 2013).

The first clues as to the potential role of snoRNAs in cancer came from scattered observations in various types of cancer, where recurrent genetic lesions seemed to target areas devoid of protein coding genes, or snoRNAs were mis-expressed in cancerous tissues compared to normal counterparts (Mannoor et al. 2012). One of the first snoRNAs identified was U50, which is implicated in prostate cancer (Dong et al. 2008) and breast cancer (Dong et al. 2009), displaying deletions, copy number losses and down-regulated expression. SNORA42 was found to be commonly overexpressed in lung cancer, suggesting an oncogenic role for the snoRNA (Mei et al. 2011) More recently, several snoRNAs have been linked to cancer, having both tumor-suppressive and pro-oncogenic roles (Mannoor et al. 2012). Moreover, the novel growth suppressor protein GRIM-1 was reported to implement its growth-inhibitory function via down-regulation of a set of box H/ACA snoRNAs, which results in disrupted rRNA maturation (Nallar et al. 2011). The association of snoRNA with disease prognosis is also documented with the unexpected finding that the snoRNAs used for normalizing miRNA expression data were actually as variable as miRNAs, and correlated with disease prognosis (Gee et al. 2011). Indeed, the applicability of snoRNA expression patterns as a prognostic tool has been suggested (Liao et al. 2010, Valleron et al. 2012b). Moreover, a multifaceted role for snoRNAs in acute leukemia was recently reported (Valleron et al. 2012a), clearly suggesting their involvement in cancer development. Interestingly, elevated expression of snoRNAs and fibrillarin have been found to be essential for tumorigenicity, and suppression of the snoRNA pathway induced p53 to mediate growth inhibition and, on the other hand, high expression of fibrillarin suppresses p53, indicating a new mechanism of nucleolar modulation of p53 (Su et al. 2013). Further, a few reports indicate that the ncRNA host genes make a snoRNA-independent contribution to cancer (Askarian-Amiri et al. 2011, Mourtada-Maarabouni et al. 2009, Williams et al. 2011). All in all, it seems that thus far only the surface has been scratched in the context of the snoRNA-cancer liaison.

RPs also play a role in tumorigenesis. Many cancer cell lines, primary tumors and leukemic cells overexpress RPs of the ribosomal subunits. The obvious reason for this would be the increased demand for functional ribosomes promoting transformation (Ruggero and Pandolfi 2003). Mutations in the RP genes RPL5 and RPL10 have been found in pediatric cases of T-cell ALL (De Keersmaecker et al. 2013).

Taken together, the relationship between cancer and the nucleolus has proved to be more complex than initially anticipated, a variety of nucleolar and cellular components being involved. Moreover, the causality would appear to be bidirectional, as mechanisms originating both from nucleolar function or from cellular requirements contribute to the emergence of cancer.

2.1.5.2 Ribosomopathies

The class of human disorders named ribosomopathies consists of inherited or somatically acquired syndromes which share a common foundation: mutations in genes encoding ribosomal components or in ribosomal assembly factors leading to impaired ribosome biogenesis. Many ribosomopathies manifest increased cancer susceptibility and impaired hematopoiesis. The ribosomopathies are summarized in Table 2.

Table 2. Summary of ribosomopathies with characteristic features.

Disease	Implicated genes	Clinical features	Of note
Diamond-Blackfan anemia	***RPS19****, RPS24, RPS17,RPL35A, RPL5, RPL11, RPS7, RPL36, RPS15, RPS27A*	Hypoplastic anemia Short stature Craniofacial defects Thumb malformation	45 % of cases familial Modest increase in cancer risk?
Shwachman-Diamond syndrome	*SBDS*	Neutropenia Anemia Pancreatic insufficiency Short stature	Autosomal recessive **Increased risk of AML**
Treacher-Collins syndrome	*TCOF1*	Craniofacial abnormalities	No hematologic abnormalities
X-linked dyskeratosis congenita	*DKC1*	Cytopenia Abnormal skin pigmentation Nail dystrophy Oral leukoplakia	**High risk of AML** **High risk for solid tumors**
Cartilage-hair hypoplasia	*RMRP*	Hypoplastic anemia Short-limbed dwarfism Hypoplastic hair Immune system dysfunction	**Increased risk of non-Hodgin lymphoma and basal cell carcinoma** Part of Finnish disease heritage
5q- syndrome	*RPS14*	Macrocytic anemia Megakaryocyte hyperplasia	**Increased risk of AML**

Modified from Liu and Ellis 2006, Narla and Ebert 2010.

One current hypothesis pinpoints ribosomal haploinsufficiency leading to disrupted ribosome biogenesis as the mechanism underlying ribosomopathies. As a result of defective ribosome biogenesis, free RPs accumulate and bind Hdm2, contributing to p53 stabilization and consequent induction of apoptosis and cell cycle arrest, as described above. Alternative mechanisms include delayed or aberrant translation of key proteins by possibly defective ribosomes, pathogenic functions of the mutated RPs or the aberrant accumulation of pre-RPs (Narla and Ebert 2010). All in all, ribosomopathies are an intriguing class of diseases, the molecular understanding of which will undoubtedly provide essential insights into basic cellular mechanisms as well as disease development and therapy.

2.1.5.3 Viral infections and the nucleolus

Viruses are intracellular pathogens which hijack cellular machineries of the host for their own replication. Viral infections have also been shown to cause changes in

nucleolar morphology, for example an increase in the size of FC and/or nucleolus, and also to cause redistribution of nucleolar proteins such as NPM and NCL. For example, during herpes simplex virus 1 (HSV-1) infection, both NPM and NCL are dispersed throughout the nucleus (Lymberopoulos et al. 2011). In addition to nucleolar localization, NCL is also present at the surface of some cells, and can act as a co-receptor for viruses upon their entry to the cell. Some viral proteins are prominent in the nucleolus, for example the regulatory Tat and Rev proteins of HIV, and NPM plays a role in their nucleolar localization. HIV utilizes the nucleolus to traffic the viral mRNA to the cytoplasm. Several viruses have been shown to interact with the nucleolus, and some exploit nucleolar proteins, for example UBF, in their replication (Greco 2009, Hiscox et al. 2010).

2.2 Leukemia

Cancer is the result of a stepwise process during which normal cells acquire several genetic and functional changes which alter their behavior and ultimately lead to malignancy. The characteristic feature of cancer cells is unrestrained proliferation, which results from the cooperation of continuous proliferative signals, evasion of signals suppressing growth or inducing cell death, and an unlimited replicative potential. For solid tumors, the ability to induce angiogenesis and to activate invasion and metastasis are advantageous attributes which further promote cancer growth and proliferation. Additional features facilitating the development and progression of cancer include increased genomic instability, reprogramming of cellular energy production, evasion of elimination by immune cells, and sustenance of tumor-promoting inflammation (Hanahan and Weinberg 2011).

Leukemia is a cancer originating in blood-forming tissue such as bone marrow, inducing the production of abnormal blood cells in large numbers, which subsequently enter the circulation and sometimes infiltrate e.g. the central nervous system (CNS). Leukemia is not a uniform disease, but rather includes a wide spectrum of diseases of the blood and/or bone marrow with a variety of clinical features and outcomes. Leukemia is characterized by defective hematopoiesis, producing abnormal blood cells which may accumulate and suppress the normal functions of the bone marrow, leading to the emergence of detectable symptoms. In leukemia, solid tumors are not formed; the malignant cells fill the bone marrow and circulate in the blood. Occasionally the cells can accumulate in the lymph

nodes or other organs. Leukemia was formerly a fatal disease with almost all patients perishing, but nowadays modern treatments result in better cure rates, approaching 90 % in some pediatric leukemias. Nevertheless, the underlying molecular pathogenesis and the mechanisms of relapse remain insufficiently understood (Bhojwani and Pui 2013, Inaba et al. 2013).

Leukemias can be classified into several subtypes. Clinically they are divided into acute or chronic types. Acute leukemia arises rapidly, whereas chronic leukemia develops gradually, and can be asymptomatic for longer periods of time. A second rough classification is based on the lineage of the cancer cells, namely myeloid or lymphoid. The malignant cells can emanate from various stages of the stem or progenitor cells differentiating into the mature cell type. In addition to these main groupings there are also more infrequent leukemias originating from other blood cell lineages.

Of all leukemias almost 90 % are diagnosed in adults. The most common forms in adults are acute myeloid leukemia (AML) and chronic lymphocytic leukemia (CLL), whereas in children the acute lymphocytic leukemia (ALL) is the most common type (85 %). Around 15 % of pediatric leukemias are AML, while chronic myeloid leukemia (CML) is very infrequent (1-2 %). The emphasis of our research was on pediatric acute leukemia, covering in fact only *de novo* acute leukemia.

2.2.1 Acute lymphoblastic leukemia (ALL)

ALL affects all age groups, but is more common in children: about 60 % of cases are diagnosed in subjects under 20 years of age. The incidence peaks in children between two and five years of age. The prognosis for pediatric ALL is good (> 85 %), whereas adults and infants face more dismal outcomes (Inaba et al. 2013). Indeed, the age at diagnosis is a strong prognostic factor: patients aged 1 - 9 years may expect a better outcome than infants or adolescents (Pui et al. 2008). This might be explained by the age-dependent pattern of occurrence of ALL subtypes. The exact pathogenetic events underlying ALL development are largely unknown, and less than 5 % of cases can be associated with predisposing genetic disorders such as Down's syndrome. The hallmark of ALL are the chromosomal translocations which often target transcription factors, thus promoting the activation of oncogenic gene expression (Pui et al. 2008). Due to advances in high

resolution analysis techniques, more detailed information regarding genetic abnormalities in ALL is available. This has led to the observation that virtually all pediatric ALL cases harbor some specific genetic anomaly. It is of utmost importance to identify such genetic abnormalities, as they are of great prognostic and therapeutic significance (Pui et al. 2011a). The targeted genes often play roles in lymphoid differentiation, tumor suppression, signaling, and regulation of cell cycle, drug responsiveness or apoptosis. There would appear to be considerable variation in the frequency of lesions among the different subtypes, as some lesions need only one or a few accompanying alterations to induce leukemogenesis (as in *MLL*-rearranged cases) whereas some ALL cases harbor several abnormalities (Pui et al. 2011a).

According to the FAB classification, ALL is divided into three subtypes: ALL-L1, ALL-L2 and ALL-L3 (see Table 1). The L2 subtype is more common in adults. Of pediatric ALL cases, about 85 % are type L1, 14 % are type L2, and 1 % type L3. However, with the exception of Burkitt's lymphoma (BLL, FAB L3), the FAB classification of ALL does not correlate well with recent immunophenotypic, cytogenetic and molecular genetic information on the ALL subtypes. This progress has led to the development of improved classification systems which incorporate all the available information, for example the WHO system. The FAB classification has nevertheless some prognostic value, with the L1 subtype associated with better outcomes than the L2 subtype, and the L3 subtype having the worst prognosis. Currently, stronger predictive factors such as age, sex, white blood cell count, treatment response (as measured by minimal residual disease, MRD) and recurrent genetic alterations have replaced the FAB classification in clinical use (Pui 2006).

The origin of ALL is in the blood progenitor cells, which are committed to the T-cell or B-cell lymphoid lineage differentiation.

2.2.1.1 Precursor B-cell ALL

The most frequent form of pediatric ALL is the precursor B-cell type. The gross chromosomal alterations most commonly found in pre-B-ALL include high hyperdiploidy with a gain of at least five chromosomes in a non-random fashion (chromosomes X, 4, 6, 10, 14, 17, 18 and 21), hypodiploidy with fewer than 44 chromosomes, recurring translocations producing fusion genes (such as *ETV6*-

RUNX1), and rearrangements of *MLL* (see Table 3) (Inaba et al. 2013). Approximately 75 % of pre-B-ALL cases show gross chromosomal rearrangements or aneuploidy. In addition, the identification of submicroscopic alterations in critical genes in over 60 % of cases makes for a more precise understanding of the genetic basis of pre-B-ALL. The critical mutations often target genes involved in B-lymphoid development (e.g. loss-of-function mutations contributing to the arrest in maturation typical of pre-B-ALL), as well as tumor suppressor and cell cycle regulatory genes leading to inactivation (Mullighan 2012).

In pre-B-ALL, the most common genetic rearrangement is the t(12;21)(p13;q22) translocation coding for ETV6-RUNX1. Both proteins are essential for normal hematopoiesis, and the fusion protein causes aberrant expression of the normal target genes. Although ETV6-RUNX1 promotes self-renewal in B-cell progenitors, it is not alone sufficient to induce leukemia. In fact, ETV6-RUNX1 is commonly detected at birth, years before the onset of leukemia, suggesting a need for secondary alterations (Bateman et al. 2010, Mori et al. 2002, Mullighan 2012). The BCR-ABL1 fusion produced by the translocation t(9;22)(q34;q11), or the Philadelphia chromosome (Ph), is present in 25 % of adult and 3 - 5 % of pediatric pre-B-ALL cases. There are two main forms of this fusion, resulting from different chromosomal breakpoints, one of which is more common in CML and the other in pediatric Ph-positive ALL. Both forms are able to activate multiple signaling pathways (Mullighan 2012). *MLL*-rearranged leukemia involves the *MLL* gene at 11q23, with over 80 identified partners. *MLL*-rearranged ALL is particularly common in infants (< 1 year), with over 90 % having the *MLL-AF4* fusion (Bueno et al. 2011). The unique characteristics of this subtype include initiation *in utero*, incorporation of both myeloid and lymphoid features, poor responsiveness to therapy, and the absence of additional genetic alterations (Mullighan 2012).

Table 3. Key genetic alterations in pediatric precursor B-cell ALL.

Subtype	Gene(s)	Frequency	Prognosis
Cytogenetic alterations			
Hyperdiploidy (>50 chr)		20 - 30%	Favorable
t(12;21)(p13;q22)	*ETV6-RUNX1*	15 - 25 %	Favorable
Trisomies 4 and 10		20 - 25 %	Favorable
t(1;19)(q23;p13)	*TCF3-PBX1*	2 - 6 %	Favorable
Intrachromosomal amp of chr 21		2 - 3 %	Poor
t(4;11)(q21;q23)	*MLL-AF4*	1 - 2 %	Poor
t(9;22)(q34;q11) or the Philadelphia chr	*BCR-ABL1*	2 - 4 %	Varied*
t(8;14)(q23;q32) or t(2;8)(q12;q24)	*MYC*	2 %	Favorable
Hypodiploidy (<44 chr)		1 - 2 %	Poor
Genetic alterations			
Overexpression of *CRLF2*	*CRLF2*	6 - 7 %	Poor
Deletion of *ERG*	*ERG*	7 %	Favorable
Cooperating mutations in *IKZF1*	*IKZF1 mut/del*	15 %	Poor
Cooperating mutations in *PAX5*	*PAX5*	30 %	?
Cooperating mutations in *JAK*	*JAK1,JAK2*	2 - 5 %	Poor

*historically dismal, outcome improved when treated with tyrosine kinase inhibitors (e.g. imatinib mesylate (Gleevec)); amp, amplification; chr, chromosome; del, deletion; mut, mutation. Modified from Mullighan 2012, Pui et al. 2011a, van der Linden et al. 2012.

2.2.1.2 T-cell ALL

Approximately 10 - 15 % of pediatric and 25 % of adult ALL are T-cell type. This is an aggressive disease, and is more common in males than females. Historically, high relapse rates are characteristic for T-ALL, and at diagnosis infiltration into the CNS is a frequent finding. Nowadays, however, the prognosis of T-ALL has improved, with 80 % of pediatric and 50 % of adult T-ALL reaching long-term remission (Van Vlierberghe and Ferrando 2012). Like pre-B-ALL, T-ALL also incorporates chromosomal translocations and aberrant expression of transcription factors (Van Vlierberghe and Ferrando 2012). Typically, a variety of oncogenic transcription factors (see Table 4) are placed under the control of potent T-cell-specific enhancers (*TCR* loci at 7q34 and 14q11), leading to their aberrant expression in T-cell progenitors (Szczepanski et al. 2010, Van Vlierberghe and

Ferrando 2012). Over 50 % of T-ALL have activating mutations in the *NOTCH1* gene encoding a transmembrane receptor normally involved in T-cell development. Additional mutations in *FBXW7*, present in about 15 % of T-ALL, further contribute to the constitutively active NOTCH1-signaling (Pui et al. 2008, Van Vlierberghe and Ferrando 2012). Deletion of the *CDKN2A* locus at 9p21, containing the *p16/INK4A* and *p14/ARF* tumor suppressor genes, is seen in over 70 % of T-ALL (Van Vlierberghe and Ferrando 2012). Several subgroups of T-ALL with specific gene expression signatures have been identified. A summary of the various factors underlying T-ALL is presented in Table 4.

Table 4. Key genetic alterations in pediatric T-ALL.

Subtype	Genes	Frequency	Prognosis
Genetic alterations			
NOTCH1 activating mutations	*NOTCH1,FBXW7*	60 %	Favorable
CDKN2A/2B deletion	*p16INK4A,p14ARF*	80 %	Favorable
Translocations to *TCR* genes	*TCR; TAL1,TAL2,LYL1, LMO1,LMO2,TLX1/HOX11, TLX3/HOX11L2,NKX2.1, NKX2.2,NKX2.5,HOX,MYC, MYB,TAN1*	40 - 50 %	Varied
TAL1 overexpression	*SIL-TAL1*	20 - 30 %	?
Fusion *NUP214-ABL1*	*NUP214-ABL1*	4 - 6%	Poor
Fusion *MLL-ENL*	*MLL-ENL*	2 - 3 %	Favorable
MLL rearrangements	*MLL*	5 - 10 %	?
Fusion *PICALM-MLLT10*	*PICALM-MLLT10*	10 %	Poor
N-RAS activating mutations	*N-RAS*	10 %	?
WT1 inactivating mutations	*WT1*	10 %	?
Other			
Early T-cell precursor phenotype		12 %	Poor

Modified from Pui et al. 2011a, Szczepanski et al. 2010, Van Vlierberghe and Ferrando 2012.

2.2.2 Acute myeloid leukemia (AML)

AML is the most common acute malignant myeloid disorder in adults, but is rare in children. The median age at AML presentation is about 70 years, and the disease is more common in men than women (Estey and Dohner 2006). Cure rates depend

on a number of prognostic factors, but generally vary between 20 - 75 %. Elderly patients tend to have poorer outcomes, partially due to their intolerance of aggressive therapies. The requirement for AML to be confirmed is that the bone marrow contain more than 30 % blast cells (Estey and Dohner 2006). In the FAB classification system AML is divided into nine subtypes (see Table 5) (Lowenberg et al. 1999).

Table 5. FAB classification of AML.

FAB type	Name	% of cases
M0	Acute myeloblastic leukemia with minimal differentiation	3 %
M1	Acute myeloblastic leukemia without maturation	15 - 20 %
M2	Acute myeloblastic leukemia with maturation	25 - 30 %
M3	Acute promyelocytic leukemia	5 - 10 %
M4	Acute myelomonocytic leukemia	20 %
M4Eo	Acute myelomonocytic leukemia with abnormal eosinophils	5 - 10 %
M5	Acute monocytic leukemia	2 - 9 %
M6	Erythroleukemia	3 - 5 %
M7	Acute megakaryocytic leukemia	3 - 12 %

Modified from Lowenberg et al. 1999.

However, with the accumulating genetic information it has been necessary to move beyond the FAB classification. The WHO system includes cytogenetic data, defines four major classes of AML, and also has a lower blast minimum criterion of 20 %. The origin of at least 20 % of the blast cells has to be of myeloid lineage, shown typically by surface markers (Estey and Dohner 2006). The most important prognostic factors are the cytogenetic and molecular genetic profiles of the malignant cells. About half of the patients present cytogenetic abnormalities at diagnosis, but the remainder with normal karyotype also harbor genetic abnormalities such as gene mutations and deregulated gene expression (Estey and Dohner 2006, Lowenberg et al. 1999). The cytogenetic features allow classification of patients into favorable, intermediate or adverse risk groups. However, there is considerable variability in treatment outcome, mostly due to the presence of secondary genetic defects (Estey and Dohner 2006). Although AML genomes harbor on average fewer mutations than most adult cancers, at least one potential driver mutation is identified in almost all cases. The genes most often mutated in AML include *FLT3*, *NPM1*, *KIT*, *CEBPA*, *TET2*, *DNMT3A* and *IDH1*, involved in e.g. transcription-factor fusions, tumor-suppression, DNA-methylation,

signaling, chromatin modifications and myeloid lineage differentiation. In addition, patterns of cooperation and mutual exclusivity imply strong biological relationships between specific alterations (Cancer Genome Atlas Research Network 2013).

AML comprises about 15 - 20 % of pediatric leukemia. Pediatric AML is a challenging disease to treat, with survival rates of about 60 %, clearly inferior to that in pediatric ALL (Radtke et al. 2009, Rubnitz 2012). Like adult AML, pediatric AML seems to carry a low burden of genetic alterations when compared to most other human malignancies. However, unique recurring genetic alterations targeting known and also potential new cancer genes have been identified, and over 90 % of cases encompass at least one genomic alteration (Pui et al. 2011a, Radtke et al. 2009). The recurrent genetic abnormalities in pediatric AML and associated prognostic significance are summarized in Table 6.

Table 6. Genetic abnormalities in pediatric AML.

Feature	Genes	Frequency	Prognosis
Rearrangements			
t(8;21)(q22;q22)	*AML1-ETO*	12 - 15 %	Favorable
inv(16)(p13;q22)	*MYH-CBF*	8 - 10 %	Favorable
t(15;17)(q22;q12)	*PML-RAR*	12 %	Favorable
t(1;22)(p13;q13)	*RBM15-MKL1*	1 - 3 %	Favorable?
t(8;16)(p11;p13)	*MYST3-CREBBP*	1 %	Poor
t(6;9)(p23;q34)	*DEK-NUP214*	1 %	Poor
11q23 *MLL* rearr.	*MLL + several partner genes*	18 - 20 %	Mostly poor
Gene mutations			
Nucleophosmin	*NPM1*	8 - 10 %	Favorable
CCAAT/enhancer BPα	*CEBPA*	4 - 6 %	Favorable
FLT3-ITD	*FLT3*	10 - 15 %	Poor
FLT3/ALM	*FLT3*	6 - 7 %	Intermediate
Wilms tumor 1	*WT1*	8 - 10 %	Poor
Rat sarcoma gene	*RAS*	5 %	?
Karyotypic changes			
del5q		1 %	Poor
-7		1 - 2 %	Poor

ALM, activating loop mutation; ITD, internal tandem duplication. Modified from Pui et al. 2011a, Rubnitz 2012.

2.2.2.1 Acute promyelocytic leukemia (APL)

APL is a distinct subtype of AML, classified as M3 in the FAB system. It is characterized by a specific translocation of chromosome 17 involving the nuclear hormone-responsive retinoic acid receptor α (*RARA*) gene. Almost all patients have the t(15;17)(q22;q21) translocation, which produces the fusion transcription factor PML-RARA, a potent repressor of nuclear receptor signaling. Also other infrequent translocation partners have been identified, including *PLZF*. Interestingly, in most cases *PML-RARA* seems to be the only driving mutation (de The et al. 2012). Typical symptoms include anemia, diminished platelet count and sudden hemorrhages due to coagulation disorders, and the disease can develop very rapidly. APL is responsive to treatment with pharmacologic doses of retinoic acid, which induces the differentiation of immature blast cells and clinical remission. The combined use of retinoic acid and cytotoxic agents usually results in a favorable outcome (de The et al. 2012, Lowenberg et al. 1999, Masetti et al. 2012).

2.2.3 Diagnosis and treatment

The classic symptoms of acute leukemia include anemia, fatigue, excessive bruising, sensitivity to infections, fever, weight loss, bone or joint pain, and enlarged lymph nodes, spleen or liver. The diagnosis of acute leukemia is made by demonstrating the accumulation of immature blast cells of over 25 %of bone marrow by e.g. May-Grünwald-Giemsa (MGG) staining. AML has to be distinguished from ALL, myelodysplastic syndromes (MDS), or AML arising in the MDS setting, as the treatments vary based on origin (Lowenberg et al. 1999). In AML, the prognostic factors are summarized, and the patient is categorized into favorable, intermediate, or adverse risk group (Estey and Dohner 2006). Most commonly, the antigens present at the blast cell surface are identified and the lineage confirmed. The exact cytogenetic and genetic features as well as the key clinical and laboratory findings are analyzed, as they are the determining factors in risk stratification. Adolescents and adults show a higher prevalence of high-risk leukemia (e.g. *BCR-ABL1*), a lower incidence of favorable subtypes (e.g. hyperdiploidy and *ETV6-RUNX1*), and a weaker tolerance of treatment, this all contributing to poorer treatment outcomes than in pre-adolescent children (Inaba et al. 2013). In ALL, the blast cells are immunophenotyped, and the subtypes with therapeutic implications, including T-

cell ALL, B-cell precursor ALL and mature B-cell (BLL) ALL, can thus be distinguished. Myeloid-associated antigen expression is detected in some 50 % of ALL cases, in distinct patterns (Inaba et al. 2013, Pui et al. 2008, Pui 2010).

The treatment of acute leukemia aims at complete remission and prevention of relapse. The latter is outstandingly important, since in most cases the true killer is the relapsed disease. The treatment consists of chemotherapy which lasts 6 - 9 months for AML and approximately 2 - 2.5 years for ALL. Usually, the approach can be divided into remission induction, consolidation (deepening the remission), and maintenance therapies. In rare cases with relapsed or resistant disease, allogeneic stem cell transplantation or irradiation may be applied (Inaba et al. 2013, Rubnitz 2012).

It is critical to monitor the response to treatment by analyzing the proportion of residual leukemic cells (MRD), by methods which allow detection below microscopical assessment, since MRD is the single most powerful prognostic factor in both pediatric and adult disease. The rate of MRD clearance in response to identical treatment varies among disease subtypes. Adults and adolescents seem to have improved outcomes with pediatric treatment compared to adult regimens (Inaba et al. 2013).

2.2.4 Developing diagnostics and treatment

Targeting the numerous variants of AML and ALL successfully remains a formidable challenge. Despite the generally good outcomes, some subtypes still carry an adverse prognosis, especially in infant and elderly patients. Significant side-effects of cytotoxic chemotherapy remain, due mainly to the systemic action of the drugs which does not discriminate between malignant and normal cells. Treatment-related short-term and long-term adverse effects should be minimized. The key is to identify the patients in need of intense therapy as against those at low risk, for whom extensive treatment is unnecessary. In part this is achieved via detailed cytogenetic, genetic and expressional analysis focusing on the leukemic cells. However, even with all this information and risk stratification, some cases behave against expectations: no response to the treatment, or an unexpected relapse, for example. Indeed, a mechanistic understanding of refractory and drug-resistant cases is often inadequate. Clearly there are additional, thus far unknown factors

which contribute to these adverse effects. The molecular mechanisms behind both *de novo* leukemia and potential relapses need to be characterized in detail to allow the development of better diagnostic tools and treatments. There is certainly a need for sensitive, accurate and cost-effective diagnostic tools. The ultimate goal is to offer every patient personalized treatment with no adverse effects, leading to permanent remission (Inaba et al. 2013, Pui et al. 2011b, Pui et al. 2012).

Novel drugs of high specificity are needed, since many of the current medications have been in use for decades, the only development being in the fine-tuning of dosing and administration schemes (Inaba et al. 2013). Many genetic lesions affecting signaling or other molecular pathways in leukemic cells have been identified, and they offer a number of potential targets for therapy. For example, the use of pharmacological inhibitors of key signaling molecules (e.g. JAKs) is being tried in cases with corresponding genetic and molecular features. The application of epigenetic treatment, including DNA methyltransferase inhibitors and histone deacetylase inhibitors, in cases with aberrant methylation status (e.g. *MLL*-rearrangements or *CREBBP* mutations) has also been suggested, as this might for instance reactivate silenced tumor-suppressor genes (Inaba et al. 2013, Pui et al. 2012). Altogether, a variety of epigenetic regulators have been found mutated in hematological diseases, this further emphasizing the role of epigenetics (Goodell and Godley 2013).

2.2.5 Leukemia heterogeneity and clonal evolution

As described, cancer is the result of a single cell lineage accumulating multiple genetic alterations. These alterations bring about genetic diversity within a cell lineage, which can result in divergent properties of the cells (subclones) e.g. to proliferate or to tolerate suboptimal conditions. In the evolutionary process of clonal progression the superior subclones are selected over the less viable and expanded (Greaves and Maley 2012). In acute leukemia the progenitor cells in the bone marrow, the self-renewing hematopoietic stem cells (HSC, HSPC), are the starting-point (Jan and Majeti 2013). There has long been ambivalence as to the nature of the mutations present in leukemic cells, i.e. which ones truly contribute to leukemogenesis. Interestingly, by sequencing two types of AML samples, Welch and associates showed that most mutations were already present in the primary HSC before the acquisition of a driver mutation, the first step in leukemogenesis

(Welch et al. 2012). Thus the originating cell may have many random but benign mutations which are "captured" in the genome and carried forward as passengers as the clone expands. In the end, the founding clone has few mutations actually driving leukemia and numerous passenger mutations with no direct contribution to leukemogenesis (Welch et al. 2012). This finding has fundamental repercussions for the way the mutational status will be perceived from now on.

The clonal evolution leading to the heterogeneity of cancer subclones is in many respects poorly understood. According to the current understanding there are several subclones present in an individual cancer. The first challenge is to determine the number and nature of these subclones, some of which are present in very low amounts. The second is to understand how the major and minor subclones, one at a time or all in concert, affect the response to treatment and the emergence of refractory disease, or contribute to relapse (Greaves and Maley 2012, Jan and Majeti 2013). Obviously, such insights will be of high clinical significance. Although there are several clonal routes to relapse, the majority of relapse cases seem to share some features with the original *de novo* leukemia. There are several hypotheses as to how relapse occurs (Jan and Majeti 2013, Mullighan et al. 2008). The simplest is that the original dominating clone (or the leukemic stem cells) survive the therapy and give rise to relapse (see A in Fig. 3). Another theory is that one of the minor subclones is resistant to therapy, persists, and is responsible for the relapse (see B in Fig. 3). An alternative is that during treatment, the DNA-damaging drugs are involved in the genetic evolution of the existing subclones, and in this way contribute to the relapse (see C in Fig. 3). A fourth option is that the DNA-damaging drugs target novel pre-leukemic cells and induce additional mutations, giving rise to a new relapsed clone (see D in Fig. 3) (Jan and Majeti 2013). Several of these mechanisms can be observed in pediatric ALL, suggesting that there are differences between patients in the relapse mechanisms or that several mechanisms act simultaneously in a single patient. It is not known to what extent treatment with cytotoxic drugs targeted to destroy the dominant clone affects disease relapse.

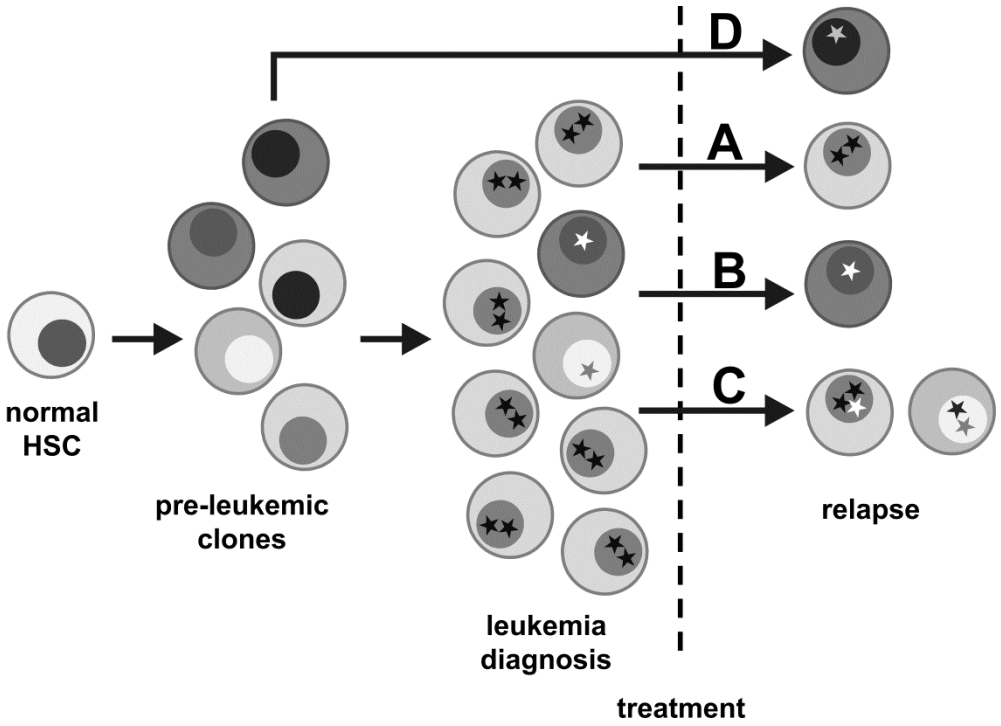

Figure 3. Clonal basis of acute leukemia evolution and relapse. A schematic illustration of the heterogeneity and clonal basis of the relationships between pre-leukemic, leukemic and relapsed cells. A normal hematopoietic stem or progenitor cell (HSC) gives rise to several normal (not shown) and pre-leukemic cells. Some of the pre-leukemic cells acquire additional mutations (indicated by a star). At diagnosis, a heterogeneous population of cells is present, and often one major subclone dominates. After treatment, there are several possible routes to relapse: A) the predominant subclone survives the treatment; B) a minor subclone present at diagnosis survives the treatment; C) the treatment induces the acquisition of additional mutation(s) in one of the subclones (major or minor) present at diagnosis; and D) the treatment contributes to the emergence of a novel leukemic clone originating in e.g. pre-leukemic cells. Figure modified from Jan and Majeti 2013, Mullighan et al. 2008.

2.3 Experimental models of leukemia

The deciphering of leukemogenesis is a multistep process. As the current understanding of leukemia is fairly advanced, the majority of the recurring genetic lesions are known. However, the difficulties lie in trying to understand how genetic changes affect the molecules, pathways and interacting mechanisms contributing to transformation. It is especially challenging to fit all the individual pieces of

information together to form an overall conception of leukemogenesis. This would not be possible without access to patient samples and information, or without experimental models to study hypotheses.

The three main resources used in leukemia studies are patient-derived material, cell lines and animal models, all of which have their advantages and disadvantages. Since due to ethical considerations it is not feasible to use human subjects in most studies, the information provided by patient samples can be further investigated by cell line and animal models.

2.3.1 Cell lines

Leukemia cell lines are immortalized cells originating from a single cell (monoclonal). Immortalization renders the cells able to evade senescence and to continuously proliferate. Human cell lines have proved rather difficult to generate, even when derived from cancerous tissues. Regardless of this, since the 1960s over 1000 human hematopoietic cell lines have been established, originating from the bone marrow or peripheral blood of ALL, AML or lymphoma patients (Drexler et al. 2000). Whereas normal hematopoietic cells can only be cultured *in vitro* for a few days, cell lines can be cultured continuously. For many leukemia subtypes, a corresponding cell line with characteristic karyotypic and cytogenetic features is available. Importantly, the cell lines retain many of the phenotypic properties and genetic alterations of the patient descent (Kennedy and Barabe 2008). Indeed, transcriptional profiling studies indicate that well characterized human leukemia-lymphoma cell lines model the gene expression pattern of the primary tumor most accurately among all tumor types (MacLeod et al. 2008). Nevertheless, it should always be kept in mind that cell lines differ in one significant aspect from the patient-derived cells: they have somehow acquired the ability to proliferate continuously. In addition, cell lines derived from other sources such as murine peripheral blood and human non-hematopoietic sources (e.g. cancer), have been utilized in leukemia research.

The major advantage of cell lines is that they provide a practically unlimited supply of homogenous material for biochemical and genetic analyses. They also make possible the worldwide availability of identical material, and can be stored viable virtually indefinitely in liquid nitrogen. Cells are quite easy to manipulate, and

various techniques to activate or inhibit e.g. gene function are widely available. However, there are also certain drawbacks. It should always be borne in mind that cell lines are not absolute replicas of the primary cancer, as both immortalization and extended culture *in vitro* can lead to the development of genetic, epigenetic and/or phenotypic changes. Also, when using cells, no systemic information is gained (Drexler et al. 2000, Kennedy and Barabe 2008). Moreover, the early events of leukemogenesis cannot be studied using fully transformed cell lines, and it is also increasingly important to aspire to experimental settings as authentic as possible in order to rule out incorrect results. For these reasons, primary human hematopoietic stem or progenitor cells have also been used in experiments (Kennedy and Barabe 2008).

Several important items of information have emerged using leukemia cell lines. Many chromosomal alterations typical of specific leukemia subtypes have been identified using cell line models. This has led to the identification of leukemia-associated oncogenes, the functional characterization of which the cell line approach has also contributed to. If necessary, cell lines can also be manipulated in various ways the better to mimic the leukemic condition. Leukemia cell lines are also useful in the development and testing of therapeutic agents. For example, the APL cell line HL-60 was used in the discovery of all-trans retinoic acid application as a differentiation therapy. More recently, leukemia cell lines have proved important in the validation of targeted molecular treatment modes such as antibody-based therapies (Kennedy and Barabe 2008) Taken together, cell lines have an established role in leukemia research, and continue to afford an important research tool.

2.3.2 Animal models

Animal models are important for an understanding of the *in vivo* and systemic events linked to leukemogenesis, especially the early stages. The mouse has proved to be a good model for various human diseases. There are gross physiological, genomic and anatomical similarities to humans, and it is relatively easy to genetically manipulate mice. Consequently, several mouse models of leukemia have been established, both by creating transgenic mice and by various manipulation techniques. Xenotransplantation, where human hematopoietic cells are introduced into another organism to study hematopoiesis and/or leukemogenesis *in vivo*, has

also been carried out using mice (Kennedy and Barabe 2008). However, species-specific differences in e.g. malignant transformation mechanisms exist, and should always be taken into account when extrapolating results to humans. In addition, in view of high costs and ethical issues, other animal models have also been established, for example the zebrafish.

2.3.2.1 The zebrafish

The zebrafish (*Danio rerio*), a vertebrate organism widely used in developmental biology, has recently emerged as a model for human cancer as well. While the zebrafish has been known to develop tumors spontaneously, the development of zebrafish lines, tools and techniques has enabled its utilization in cancer research. Zebrafish offer several advantages compared to mice. For example, maintenance cost and effort are lower, offspring production is efficient (hundreds of fertilized eggs per pair of adult fish in a week), the embryos and larvae develop externally and are optically transparent for up to 30 days. Also, the complete zebrafish genome is available, and many human genes have conserved counterparts in the zebrafish, the protein products of which show structural and functional similarity to the human. The development and progression of cancer in the zebrafish is similar to that in humans in many key respects. Thus far, a major advantage of the zebrafish is the remarkable transparency during development and also in adulthood when using a transparent mutant fish line, which allows *in vivo* imaging of many cancer-associated processes. In addition, zebrafish successfully host xenografts such as human cells, efficiently take in small molecular compounds directly from water, and are amenable to large-scale screening due to cost-effectiveness and availability in large supply. All these attributes make zebrafish an attractive cancer model, which has been applied in e.g. analyses of chemical mutagenesis, screening for anti-cancer agents, creation of cancer mutant and transgenic fish lines, and xenotransplantation of human cancer cells to investigate tumor cell behavior *in vivo* (Stoletov and Klemke 2008).

Zebrafish hematopoiesis is markedly similar to that of humans at the genetic and cell biological level. Even though hematopoiesis takes place in the kidney marrow of zebrafish, and there are some differences in the appearance of the cells (e.g. zebrafish erythrocytes are nucleated), the major molecular pathways are conserved. The zebrafish has been thus found suitable to study hematological

disorders and leukemia. The presence of aberrant fusion genes is characteristic of acute leukemias. The zebrafish offers several possibilities to study the function of these fusion genes and the fusion proteins they produce *in vivo*. Transient or persistent expression can be achieved by microinjecting synthetic mRNA or a plasmid construct into zebrafish embryos. The expression of genes of interest can also be abrogated transiently (antisense morpholino oligonucleotides) or permanently (zinc-finger-mediated endonuclease technology). Similarly, suspected oncogenes and tumor suppressor genes can be functionally characterized using these methods, and completely novel gene functions can be identified via extensive forward genetic screens. In addition, the xenograft approach to screen systemic drug effects and efficacy has been applied in the zebrafish. It is also possible to follow closely the different phases in the development of the disease. It follows that a number of transgenic zebrafish leukemia models have been developed. Indeed, transgenic fish have been shown to develop leukemia with similar morphology and genetic background to human leukemia, with sensitivity to the same chemotherapeutics. Moreover, the transgenic leukemic fish lines can be used to screen for suppressors and enhancers of leukemogenesis. The major disadvantage of using transgenic zebrafish as models of leukemia has thus far been the low-penetrance long latency of the disease (Lohi et al. 2013, Payne and Look 2009, Teittinen et al. 2012).

3. Aims of the study

The aim of this series was to shed light on the potential role of the nucleolus in pediatric acute leukemias. Two major components of the nucleolus, the small nucleolar RNAs (snoRNAs) and the nucleolar proteins, were studied in various leukemia subtypes. Differences in expression patterns were analyzed using cell line models and patient samples. In addition, a model system to study the *in vivo* function of nucleolar proteins was set up using a known nucleolar protein as pioneer.

The specific aims were:

1. to study the expression of snoRNAs in leukemic cell lines representing major leukemia subtypes.

2. to study the expression of snoRNAs in normal and patient samples.

3. to identify nucleolar proteins whose expression is altered in leukemic cells.

4. to set up a system to study the function of nucleolar proteins in the context of a living organism, the zebrafish.

4. Materials and methods

4.1 Cell culture (I, III)

The cell lines used in these studies are listed in Table 7. Leukemic cell lines were obtained from the Leibniz Institut DSMZ (German Collection of Microorganisms and Cell Culture, Braunschweig, Germany). KOPN-8, Kasumi-2, Nalm-6, REH, CCRF-CEM, Molt-16, Jurkat, P12-Ichikawa, HL-60 and Daudi cells were cultured in RPMI-1640 medium (Gibco, Invitrogen) supplemented with 1 % L-glutamine (L-glu), 10 % fetal bovine serum (FBS), and 1 % penicillin/streptomycin antibiotics (pen-strep). MHH-CALL3 and Peer cells were cultured in the same medium containing 20 % FBS. OCI-AML3 cells were cultured in MEM Alpha Medium (Gibco, Invitrogen) supplemented with 1 % L-glu, 20 % FBS and 1 % pen-strep. HeLa cells were cultured in D-MEM medium (Sigma-Aldrich) supplemented with 1 % L-glu, non-essential amino acids, 10 % FBS and 1 % pen-strep. Cell culturing was carried out as instructed by DSMZ.

Table 7. Cell line characteristics.

Cell line	Type*	Karyotype, fusion genes*	Age of donor*	Study
CCRF-CEM	T-cell ALL	near tetraploid	3 years	I, III
DAUDI	BLL	near diploid	16 years	I
HeLa	cervix carcinoma	hypertriploid/hypotetraploid	31 years	I, III
HL-60	AML-M2	hypotetraploid/hypodiploid	35 years	I, III
JURKAT	T-cell ALL	hypotetraploid	14 years	I
KASUMI-2	pre-B ALL	hypodiploid	15 years	I
KOPN-8	pre-B ALL	hypodiploid, MLL-ENL	3 months	I, III
MHH-CALL3	pre-B ALL	pseudodiploid	11 years	I, III
MOLT-16	T-cell ALL	near diploid	5 years	I, III
NALM-6	pre-B ALL	near diploid	19 years	I
OCI-AML3	AML-M4	hyperdiploid, NPM mutation	57 years	I, III
P12-Ichikawa	T-cell ALL	hypotetraploid	7 years	I
PEER	T-cell ALL	pseudodiploid, NUP214-ABL1	4 years	I
REH	pre-B ALL	pseudodiploid, TEL-AML1	15 years	I

*data from DSMZ

4.2 RNA extraction and quantitative real-time PCR

4.2.1 Extraction of RNA (I, IV)

The mirVana miRNA isolation kit (Ambion, Applied Biosystems) was used to extract the small RNA fraction (less than 200 nucleotides in length) from the cultured cells (80 - 90 % confluency) following the manufacturer's instructions (Study I). In the case of zebrafish embryos and adult fish tissues (study IV), the total RNA was extracted with an RNeasy Mini Kit (Qiagen) as instructed by the manufacturer. The quality and concentration of the RNA extracts were checked using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

4.2.2 Quantitative real-time PCR (I, IV)

For quantitative real-time PCR (RT-qPCR), the RNA was converted to cDNA template with an iScript Select cDNA Synthesis Kit (Bio-Rad) and random primers according to the manufacturers' instructions. RT-qPCR was performed using the Evagreen Ssofast supermix kit (Bio-Rad) and a CFX96 real-time thermal cycler (Bio-Rad) as instructed by the manufacturer. Primer sequences are given in Table 8. For snoRNA expression in Study I, the U6 RNA was used for normalization of data (Gee et al. 2011, Schmittgen et al. 2004). In Study IV, the expression of the *SAP30L*, *nkx2.5* and *ppox* genes in the zebrafish was normalized to the housekeeping gene *EF1A* (Tang et al. 2007). An average threshold cycle (CT) value was calculated from two replicate samples, and quantitation performed using the comparative CT method (Schmittgen and Livak 2008). The runs were repeated twice or thrice, and no significant differences between replicate runs were observed. When appropriate, differences between groups were compared using the two-tailed Student t test.

Table 8. Sequences of RT-qPCR primers.

Target	Primers (5' → 3')	Study
EF1A (dr)*	CTG GAG GCC AGC TCA AAC AT ATC AAG AAG AGT AGT ACC GCT AGC ATT AC	IV
nkx2.5 (dr) **	GCG AAG ACC TTC AGG AGG ACA AAG GCA AC CGA GGC TTC CTC CTC TTC CTC TGC TTG GG	IV
ppox (dr)	GTG ATC AGT AAA GGG ACA GAA CTG CAT GCC CCT CTT TAG	IV
SAP30L (dr)	AAC GCG TCC TTC AGC AAG AGG AT GTT GAC CTG AAG CTG GAA CAG ATC	IV
scaRNA6	AGG TCG ATG ATG ATT GGT AAA AGG AGG TCT CAG ATT GAA AAC TTG AG	I
scaRNA9	CTG CTT TTA GTG AAG TGG AT CCA GAC ATA TGC CCT TAT T	I
SNORD24	TGC AGA TGA TGT AAA AGA A TGC ATC AGC GAT CTT GGT	I
SNORD44	CCT GGA TGA TGA TAA GCA A TTA GTC AGT TAG AGC TAA T	I
SNORD49A	TGC TCT GAT GAA ATC ACT AAT CAG ACA GGA GTA GTC TT	I
SNORD55	GTG TAT GAT GAC AAC TCG G GCT CAG CTC TCC AAG GTT G	I
SNORD82	CAG CAC AAA TGA TGA ATA AC AGC ACA TCA GCA CAC TAC AAT	I
SNORD105	CCC CTA TCT CTC ATG ATG CCC CAT CTC TTC TTC AGA	I
SNORD110	CAG TGA TGA CTT GCG AAT C GAG ACA TGG AGA CAT CAG TGA	I
U6 ***	CTC GCT TCG GCA GCA CA AAC GCT TCA CGA ATT TGC GT	I

dr, *Danio rerio*; * primers from Tang et al. 2007; ** primers from Targoff et al. 2008; *** primers from Schmittgen et al. 2004.

4.3 Microarrays, massive parallel sequencing and data analyses

4.3.1 SOLiD sequencing and data analysis (I)

The snoRNA transcriptome was analyzed from two independent small RNA extractions per cell line. For each sample, the fragment library for SOLiD sequencing was prepared following the SOLiD small RNA expression kit protocol provided by the manufacturer. The samples were processed on a SOLiD 4

instrument (Applied Biosystems, Life Technologies) as instructed. The read length was 35 bp, and two independent RNA samples were analyzed from each cell line. The human reference genome (hg18) and the SOLiD BioScope v.1.3 software package were used to map the data. The Functional RNA Database, a non-coding RNA sequence database (http://www.ncrna.org/frnadb/) v. 3.4, was used to retrieve snoRNA features against which the tags were counted. The R/Bioconductor software was used for downstream analysis, and the DESeq analysis package (Anders and Huber 2010) for normalization. The Limma package (Smyth 2005) was used for statistical testing, and features with fold change $FC \geq |1.5|$ and p-value $p<0.05$ were considered to be differentially expressed. Heatmap clustering of the data was done with the GENE-E tool (http://www.broadinstitute.org/ cancer/software/GENE-E/).

4.3.2 Hematological gene expression dataset (II)

The compilation of the hematological gene expression dataset containing both neoplastic and non-neoplastic samples is described in detail by Heinäniemi and associates (Heinaniemi et al. 2013). Briefly, the published gene expression datasets obtained using HG-U133 Plus 2.0 GeneChip platform (Affymetrix) were sought and downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/), a public functional genomics data repository. To obtain an acceptable level of comparability between the samples, quantile normalization and median polishing were performed on the raw probe values in a MATLAB environment. Expression levels for genes were summarized from probes which aligned uniquely to the corresponding transcripts as listed in the Reference Sequence (RefSeq) database. Only genes with at least four uniquely aligning probes were selected so as to filter out unreliable expression values. Fifteen genes on the microarray were identified to be snoRNAs.

4.3.3 Microarray and data analysis (IV)

The samples were prepared from zebrafish total RNA as recommended by the manufacturer (Agilent Technologies). Agilent's 4x44K Zebrafish V3 Gene Expression Microarray was used. The array was scanned with an Agilent Technologies scanner (model G2565CA), using scan profile

AgilentHD_GX_1Color. Numerical data were extracted with Agilent's Feature Extraction software, version 10.7.1., using grid 026437_D_F_20100719 and protocol GE1_107_Sep09. The R programming language and environment (R Development Core Team 2008) and its Bioconductor module (Gentleman et al. 2004) were used to analyze the microarray data. Quantile normalization was used, and the quality of the data was checked by several quality control methods. Statistical testing was performed with the Limma package (Smyth 2005), and stringent filtering thresholds were used (fold change $FC \geq |3|$ and p-value $p \leq 0.001$).

4.4 Isolation of nucleoli (III)

For the isolation of nucleoli, a cell fractionation protocol based on sucrose gradients by Andersen and colleagues (Andersen et al. 2002) was used with some modifications. The cells were grown to approximately 90 % confluency and 1-3 x 10^8 cells collected by centrifugation for each isolation. The purified nucleoli were stored at -70 °C. The total protein concentration of the nucleolar extracts was determined using a standard protein assay (DC Protein Assay, Bio-Rad) following the manufacturer's protocol.

4.5 Two-dimensional difference gel electrophoresis (2-D DIGE)

4.5.1 2-D DIGE (III)

2-D DIGE has two dimensions: first, the isoelectric focusing (IEF), which separates proteins by their isoelectric point (pI), and second, SDS-PAGE, which separates proteins based on molecular weight. For 2-D DIGE, the Ettan DIGE system (GE Healthcare) was used, following the manufacturer's instructions. The nucleolar samples were from CCRF-CEM, MHH-CALL3, OCI-AML3 or HeLa cell lines. The samples were labeled with fluorescent cyanide minimal dyes (CyDye DIGE Fluor dyes, GE Healthcare) according to the manufacturer's protocol. A combined mix of two samples (labeled with CyDye3 or CyDye5) and an internal standard (labeled with CyDye2), with 150 μg of protein in total, was applied per 2-

D DIGE analysis. IEF was performed with non-linear immobilized pH gradient strips using an Ettan IPGphor II Isoelectric Focusing System (GE Healthcare). Subsequently, the IEF-separated samples were run on 12 % SDS-PAGE gels using the Ettan DALTsix system (GE Healthcare). A schematic illustration of the two dimensions of 2-D DIGE is given in Figure 4.

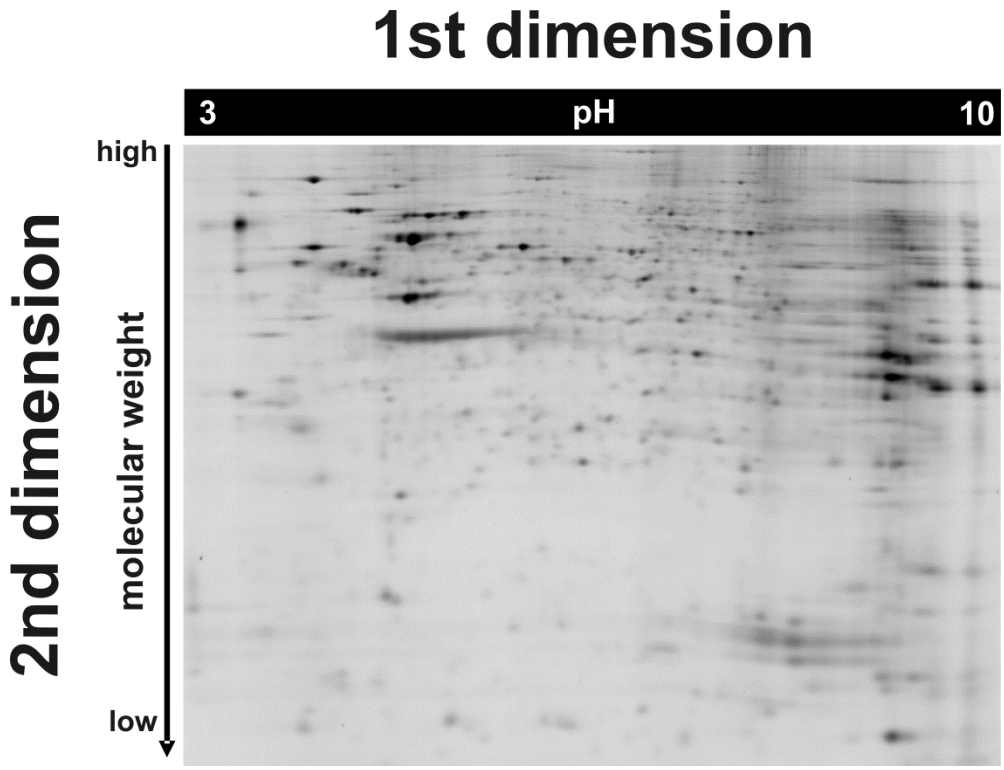

Figure 4. 2-D DIGE. A schematic illustration of the two dimensions of 2-D DIGE. In the first dimension, isoelectric focusing (IEF), the proteins in the sample are separated by their isoelectric point (pI) in pH gradient. In the second dimension, the proteins are separated based on their molecular weight in SDS-PAGE. Thus each spot in the gel contains proteins of similar size and pI. A silver-stained 2-D DIGE gel is shown in the figure.

4.5.2 Gel imaging and protein identification (III)

CyDye-labeled protein spots were visualized with a TyphoonTM Trio Variable Mode Imager (GE Healthcare) at the appropriate wavelengths. The spot map

images were analyzed using DeCyderTM v6.5 software (GE Healthcare) according to the manufacturer's recommendations, and p-values were calculated using Student's test. The gels were post-stained with silver using a method compatible with identification of proteins by mass-spectrometry (MS) (Yan et al. 2000), and spots of interest were cut and processed for MS analysis. An automated nanoLC-MS/MS consisting of aQSTAR® Elite Hybrid LC/MS/MS System (Applied Biosystems) coupled to a nano-LC instrument UltiMate® 3000 (Dionex Corporation) was used. Database searches with the MS/MS data obtained were performed with the MASCOT MS/MS Ion Search program (http://www.matrixscience.com) against the MSDB database.

4.6 Western blotting and staining

4.6.1 Western blotting (III, IV)

The cells were lysed in RIPA lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl 1 % Igepal (NP-40), 0.5 % Na-deoxycholate, 0.1 % SDS) supplemented with Complete mini protease inhibitor cocktail (Roche). In study III, the total protein concentration of the samples was determined using a standard protein assay (DC Protein Assay, Bio-Rad), and samples with uniform protein concentrations were prepared in 2 x Laemmli buffer. In study IV, 2 dpf zebrafish larvae were lysed directly in 2 x Laemmli buffer (10 larvae/ 50 µl), by boiling and vigorous pipetting, and the cell debris was removed by centrifugation. Proteins were resolved by SDS-PAGE (Mini-PROTEAN Tetra Cell system, Bio-Rad) and transferred to nitrocellulose membranes (Amersham Biosciences). The immunoblots were blocked with 5 % non-fat milk, probed with the desired primary antibody overnight, and after washing probed with a suitable HRP-conjugated secondary antibody (DakoCytomation). Enhanced chemiluminescence was used in detection. Primary antibodies used in Western blotting and immunofluorescence staining are summarized in Table 9.

4.6.2 Immunofluorescence staining (III)

The presence of the protein of interest in nucleolar extracts was analyzed by immunofluorescence staining. Smear preparations of the extracts were prepared on Superfrost Plus microscope slides (Thermo Scientific). The cells were fixed with 4 % PFA-PBS for 20 minutes, rinsed with PBS, and permeabilized in 0.2 % solution of Triton X-100 in PBS for 10 minutes. Unspecific binding of the antibodies was blocked by incubation in 1 % BSA-PBS at 37 °C for 45 minutes. Incubation in the primary antibody (see Table 9) solution was carried out at 37 °C for 1 hour, followed by washes in PBS. AlexaFluor-conjugated secondary antibodies (Molecular Probes, Invitrogen) were used (1 hour at 37° C), and, after washes in PBS, the cells were mounted on a DAPI mount (VectaShield, Vector Laboratories). Stainings were analyzed and photographed using a BX60 microscope and the Cell^D software (Olympus).

Table 9. Antibodies used in Western Blotting and immunofluorescence staining.

	Dilution used in			
Antibody	**WB**	**IF**	**Product #**	**Manufacturer**
Nucleophosmin (NPM)	1:8000	1:2000	32-5200	Invitrogen
Prohibitin (PHB)	1:1000	1:50	sc-28259	Santa Cruz Biotechnologies
SAP30L	1:250	-	-	see Korkeamäki et al. 2008
TDP-43	1:1000	1:150	T1705	Sigma-Aldrich
TDP-43-N	1:1000	1:150	SAB4200006	Sigma-Aldrich

IF, immunofluorescence; WB, Western blotting.

4.6.3 Histological staining (IV)

The morphant and control zebrafish larvae were fixed (4 % PFA-PBS, 4 °C), and embedded in 2 % agarose. Dehydration was performed by incubation in a series of alcohol solutions (70 %, 96 %, and absolute ethanol; 1 - 2 h in each), and last in xylene for at least 1 h. Next, the samples were embedded in paraffin and 5 μm sections were cut in longitudinal orientation. The sections were fixed on glass slides and deparaffinized by incubation in xylene (2 x 4 min), absolute ethanol (2 x 3 min), 96 % ethanol (2 x 3 min), once in 70 % ethanol, and rinsing once with water. Hematoxylin-eosin staining was performed by incubating the slides in Mayer's hematoxylin (2 min), running water (2 min), water (1.5 min), 70 % ethanol (15 s), eosin Y (15 s), 96 % ethanol (2 x 30 s), absolute ethanol (2 x 1 min), and xylene (1-

4 min). Thereafter the slides were mounted and pictures taken under a BX60 microscope using the Cell^D software (Olympus).

4.7 Zebrafish methods (IV)

4.7.1 Zebrafish maintenance

Wild type AB zebrafish were maintained under standard conditions at 28.5 °C as described (Westerfield 1995) at the Tampere Zebrafish core facility. The care of the animals was in accordance with the Finnish Laboratory Animal Welfare Act 62/2006 and the Laboratory Animal Welfare Ordinance 36/2006.

4.7.2 Whole-mount *in situ* hybridization (WISH)

The sense and antisense probes for zebrafish SAP30L were prepared using the DIG RNA Labeling Mix (Roche Diagnostics) and linearized plasmid templates (containing full-length SAP30L cDNA in sense or antisense orientation under T7 promoter) according to the manufacturer's instructions. Zebrafish larvae were collected at 12, 24, 36, 48, 72 and 120 hpf, placed immediately in 4 % PFA-PBS solution, and fixed overnight at 4 °C with gentle agitation. The WISH protocol by Thisse and Thisse (Thisse and Thisse 2008) was followed with some modifications. Briefly, the fixed embryos/larvae were rehydrated in 2 ml tubes, with incubations in a descending methanol-PBST series, and washed in PBST. Digestion with proteinase K (10 μg/ml in PBS) was done at 4 °C for 20 s to 40 min according to developmental stage, followed by washing in PBST, refixing in 4 % PFA-PBS, and subsequent washing in PBST. The larvae were prehybridized for 2 - 3 h at 65 °C, followed by overnight incubation in the hybridization mix containing 130 ng of DIG-labeled RNA probe, (SAP30L antisense or sense control), at 65 °C. Next, the larvae were washed in a series of pre-hybridization-SCCT-PBST solutions, first at 65 °C, then at room temperature. Thereafter the larvae were blocked for 1 - 2 h, and subsequently incubated in the anti-DIG antibody (1:2000 dilution, sheep anti-digoxigenin-AP Fab fragments, Roche Diagnostics) overnight at 4 °C. The larvae were washed, and incubated in the staining solution (premade BM Purple AP substrate, Roche) until desired staining intensity was reached. The reaction was

stopped, and the staining was analyzed and photographed using a dissecting microscope. The anatomical reference images by Haffter and colleagues (Haffter et al. 1996) were used to assess the localization of the staining.

4.7.3 Morpholino knock-down and mRNA rescue

In microinjection experiments, embryos at the 1 - 4 cell stage were used. The injection was done into the yolk sac with 1 nl of injection mix with the PV830 Pneumatic PicoPump injection system (WPI, Inc.). Rhodamin dextran was included in the injection mix to allow selection for successfully injected embryos. The translational start site of the zebrafish SAP30L mRNA (Ensembl ID: ENSDARG00000030213) was targeted with two independent antisense morpholino oligonucleotides: SAP30L-MO1 (5' CCT CCT CCG TGC TGA AGC CGT TCA T) and SAP30L-MO2 (TCA TCC CGG AGA CCG GAC ACC GAG C). The random control morpholino (RC) and the p53-targeting morpholino (Robu et al. 2007) were also used. All morpholinos were purchased from Gene-Tools, LLC. The efficiency of SAP30L-MO1 in blocking the SAP30L mRNA translation was confirmed using an *in vitro* transcription and translation assay (TNT kit, Promega). The effective concentrations of morpholinos were experimentally determined. In the mRNA rescue experiments, SAP30L-MO1 or SAP30L-MO2 and a capped synthetic mRNA encoding full-length zebrafish SAP30L or full-length zebrafish nkx2.5 were simultaneously injected. The mRNA was synthesized using a T7 mMESSAGE mMACHINE Kit (Ambion) and corresponding plasmid templates according to the manufacturer's instructions.

4.7.4 Analysis of heart morphology and function

Heart morphology was assessed and heart rates determined visually in live zebrafish morphant and control larvae at 5 dpf. The longitudinal fractional shortening of the atrium and ventricle were calculated according to the formula (LDd - LDs)/LDd, where LDd is the longitudinal diameter at diastole and LDs the diameter at systole.

4.7.5 *o*-dianisidine staining

Detection of hemoglobin by *o*-dianisidine staining was performed as previously described (Ransom et al. 1996). The zebrafish larvae were stained at 3 dpf, analyzed under a dissecting microscope, and scored as normal, reduced, or severely reduced based on the qualitative level of staining, which was assessed visually.

4.8 Statistical considerations (IV)

Quantitative data are given as mean ± standard deviation. When appropriate, differences between groups were compared using the two-tailed Student t test or the $\chi 2$ test for cross tabulations. A p-value of less than 0.05 was considered statistically significant. Statistical calculations were performed using PASW Statistics version 18.

4.9 Ethical considerations (I – IV)

In this series, commercial cell lines were used. The patient data included in the hematological gene expression dataset have been collected from published datasets. The zebrafish work was carried out at the Tampere Zebrafish core facility where the zebrafish stocks are kept and raised under permission of the State Provincial Office of Western Finland (permission LSLH-2007-7254-/Ym-23). Since zebrafish embryos and larvae up to the age of seven days were used in this study, a separate permission is not required (the effective practice in Finland). The zebrafish can be considered as the most ethical choice for this study as it is the least neurophysiologically developed genetically tractable model organism.

5. Results

5.1 Small nucleolar RNAs in leukemia

5.1.1 The expression of snoRNAs in leukemic cells (I)

Many recent reports demonstrate the involvement of ncRNAs in various normal as well as pathogenic processes in the cell, expanding our understanding of these multifaceted molecules far beyond the previously recognized roles. Although snoRNAs have a well-established role in guiding rRNA modifications, sporadic observations on their participation in additional functions have accumulated (Mannoor et al. 2012). Prompted by these observations and the fact that the characteristics of nucleoli vary in ALL blasts, we hypothesized that the expression of snoRNAs may be altered in leukemic cells. To address this, we chose two leukemic cell lines representing each of the main acute leukemia subtypes (T-ALL, pre-B-ALL, and AML), and one cell line representing BLL (see Table 1 in Study I). Also, mononuclear cells (MNC) and HeLa cells were included in the analysis in order to provide a normal as well as non-leukemic cancer point of comparison, respectively. The small RNA fractions were isolated from the cell lines (two independent isolations per line), and the small RNA transcriptome was analyzed by massive parallel sequencing (SOLiD 4, Life Technologies). The snoRNAs were identified, and the normalized expression values for all snoRNAs in each sample were determined. This snoRNA expression pattern or "snoRNA fingerprint" is unique for each sample. The relationships between samples can be studied using cluster analysis, where the samples are grouped according to their general similarity. Interestingly, the individual cell line samples clustered together according to leukemia subtype (see Fig. 5). This analysis thus implies that snoRNAs are differentially expressed in different leukemia subtypes. It is noteworthy that normal MNCs differ significantly from both HeLa and leukemic cells in their snoRNA expression pattern (data not shown).

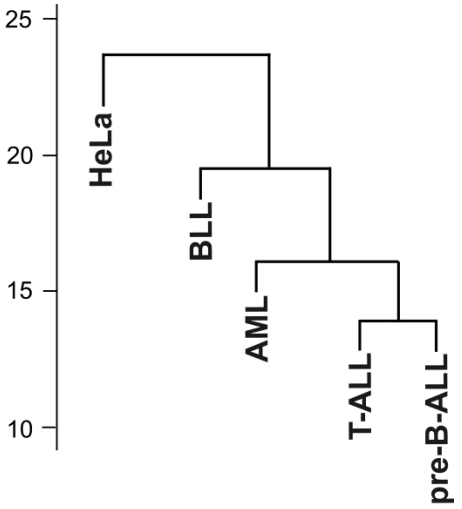

Figure 5. Hierarchical clustering of samples based on normalized snoRNA expression. When hierarchical clustering is visualized by a dendrogram, most similar samples are found in branches closest to one another. Figure modified from Study I.

5.1.2 Differential expression of snoRNAs in leukemia subtypes (I)

Next, we filtered out the differentially expressed (DE) snoRNAs between AML, T-ALL and pre-B-ALL subtypes. The filtering was done based on fold change (FC) and p-values with cut-offs set at FC >|1.5| and $p<0.05$. A heat map clustering of the DE snoRNAs by leukemia subtype is shown in Fig. 2 in Study I. Several DE snoRNAs were identified to be up- or down-regulated in the comparisons between AML and ALL (both subtypes) and between T-ALL and pre-B-ALL, and are summarized in Table 10. Of the snoRNAs (including scaRNAs) listed in the snoRNA-LBME-db (Lestrade and Weber 2006), approximately 67 % are box C/D, 27 % are box H/ACA snoRNAs, and roughly 6 % are scaRNAs. Interestingly, the majority of the DE snoRNAs were box C/D type. Furthermore, several scaRNAs were also differentially expressed.

Table 10. Differentially expressed snoRNAs in the comparisons.

Comparison	DE snoRNAs	up-regulated	types	down-regulated	types
AML vs. ALL	**40**	**22**	box C/D: 17 box H/ACA: 3 scaRNA: 2	**18**	box C/D: 18 box H/ACA: 0 scaRNA: 0
T-ALL vs pre-B-ALL	**46**	**29**	box C/D: 19 box H/ACA: 5 scaRNA: 5	**17**	box C/D: 15 box H/ACA: 1 scaRNA: 1

As most pediatric leukemias are ALL, we next focused on a comparison between the two main classes of ALL, pre-B-ALL and T-ALL. In the comparison between T-ALL and pre-B-ALL, 65 % of the up-regulated snoRNAs were box C/D, and box H/ACA and scaRNAs accounted for 17 % each. Among the down-regulated, box C/D snoRNAs dominated with 88 %, whereas both box H/ACA snoRNAs and scaRNAs stood at 6 %. Seven snoRNAs and two scaRNAs up- or down-regulated in this comparison were selected (see Table 2 in Study I, and Table 11), and their expression was analyzed by RT-qPCR. In the first round of RT-qPCR, the pre-B-ALL and T-ALL cell lines used in SOLiD analysis were included, and eight out of the nine snoRNAs showed changes in expression similar to those observed in the SOLiD analysis. In the second round of RT-qPCR, six additional cell lines were included in the analysis, and similar changes in expression were seen in six out of the nine snoRNAs (see Fig. 3 in Study I). The results of the analyses of differential expression between T-ALL and pre-B-ALL for the nine snoRNAs are summarized in Table 11. We also tested how the total RNA fraction not enriched for small RNA performed as template in the analysis. The results were weakly repeated, and for some snoRNAs the expression levels were very low, which renders interpretation of the result unreliable.

Table 11. Summary of the expression analyses of the nine snoRNAs from the T-ALL versus pre-B-ALL comparison. The difference in expression revealed by SOLiD sequencing analysis is given as fold change (FC) value. RT-qPCR was used to analyze the change in expression between the two ALL subtypes, and is indicated relative to the SOLiD FC (similar or opposite). In Round 1, small RNA templates from 2 T-ALL and 2 pre-B-ALL cell lines were analyzed, and Round 2, small RNA templates from 5 T-ALL and 5 pre-B-ALL cell lines. The cell lines are given in Table 1 in Study I.

	SOLiD analysis	Change in expression by RT-qPCR	
name	**FC**	**Round 1**	**Round 2**
scaRNA6	3.284	similar	similar
scaRNA9	1.938	similar	similar*
SNORD24	-1.485	similar	similar*
SNORD44	-1.826	similar	similar
SNORD49A	1.965	similar	opposite
SNORD55	1.723	similar	opposite
SNORD82	-1.531	similar	similar
SNORD105	1.513	similar	no change
SNORD110	-1.737	opposite	similar

* $p<0.05$; FC, fold change

5.1.3 Expression of snoRNAs in hematological samples (II)

Prompted by the observation that snoRNAs are differentially expressed in cell lines representing acute leukemia subtypes, we sought next to study whether this is also the case in normal hematopoietic and leukemic cells originating in patients. To this end, we performed an *in silico* analysis by utilizing a unique hematological database with a large cohort of pediatric sample data produced using an Affymetrix Microarray platform (Heinaniemi et al. 2013). The 15 snoRNAs present on this microarray and their characteristics are listed in Table 1 in Study II. The expression dataset contains 883 cases of pediatric acute leukemia representing the main subtypes AML (237 samples), T-ALL (118 samples), and pre-B-ALL (or early-B-ALL, 528 samples). Furthermore, there are also non-neoplastic hematopoietic samples available in the expression dataset. Altogether 35 such samples were included in our analysis, among them seven HSC, five naïve B-cell, 10 naïve $CD4^+$ T-cell, and 13 $CD8^+$ T-cell samples.

To begin with, the relative expression levels in different leukemia and normal blood cell subtypes were compared. The overall expression levels of all snoRNAs appeared to be relatively uniform in both normal and leukemia samples, as in most cases no substantial differences in the expression levels between the cell subtypes could be discerned (see Fig. 1a in Study II). Two snoRNAs, SNORA70 and SNORD104, were expressed at a higher level compared to the others. The approximately fourfold higher level of expression was seen in both healthy and leukemic samples. In addition, the expression of SNORA25 and SNORA61 was strong in naïve B-cells compared to the other normal or leukemia cell types (Fig. 1a in Study II), suggesting that the expression of these snoRNAs is regulated in a manner dependent on B-cell differentiation. When the individual snoRNAs were examined at the level of single patients, fairly uniform expression was yet again observed. The only distinguishable exception was SNORD114-3, which is discussed below.

SNORD114-3 expression in the patient samples was fairly consistent except for 15 cases among the AML samples, in which the expression was increased (threshold level set at 7 in log2-expression). When these 15 cases were examined more closely, it was observed that 13 of them carried the t(15;17) translocation, a feature distinctive of acute promyelocytic leukemia (APL). Altogether 19 pediatric APL cases were included in the samples, the rest of which displayed SNORD114-3 log2-expression under 7 (Fig. 1b in Study II). When the pediatric AML samples are arranged based on SNORD114-3 expression level, the APL cases are clearly enriched among the high-expressing samples. The level of SNORD114-3 expression also differentiates between APL cases and other AML cases in a statistically significant manner (Fig. 1c in Study II). Furthermore, when the association between high SNORD114-3 expression and APL was studied in the 1117 adult AML samples found in the expression dataset, a similar phenomenon was seen (Fig. 1c in Study II). Thus, high expression of SNORD114-3 denotes APL. This has also been reported elsewhere (Cohen et al. 2012, Valleron et al. 2012a).

5.1.3.1 The imprinted locus DLK1-DIO3 in leukemia (II)

SNORD114-3 is located in the chromosomal region 14q32, and is one of the 31 snoRNAs in the SNORD114 cluster in the *DLK1-DIO3* domain. Interestingly, this

domain is genomically imprinted, which is an epigenetic mechanism leading to mono-allelic gene expression in a parent-of-origin specific manner. Over 100 imprinted genes have been discovered in humans, many of which operate during development, and are usually organized in conserved clusters regulated by an imprinting control region (ICR). ICRs are long elements rich in CpG elements, which are heavily methylated only on the paternally- or maternally-inherited allele, repressing expression from this allele. Thus, genes in these regions are particularly susceptible to errors, and many diseases such as cancer and developmental syndromes are associated with disturbances in imprinted domains. Especially interesting is that 117 snoRNAs have been found to be imprinted, and most of them are clustered in large transcribed units which produce many highly similar snoRNAs (Girardot et al. 2012).

The imprinted *DLK1-DIO3* domain contains two box C/D snoRNA clusters, several miRNAs, and protein-coding and non-protein-coding transcripts (see Fig. 6). The three protein-coding genes (*DLK1*, *RTL1*, and *DIO3*) are expressed from the paternal allele, whereas the maternal allele gives rise to two non-protein coding genes (*MEG3* and *MEG8*, or RTL1-antisense), as well as SNORD112, SNORD113 cluster of 9 snoRNAs, SNORD114 cluster of 31 snoRNAs, and in total 53 miRNAs (Girardot et al. 2012). Interestingly, the snoRNAs in this domain are orphans, that is, they lack a known target. In order to establish whether other transcripts originating in this locus show similar deregulation as SNORD114-3, their expression was analyzed. Indeed, *DLK1*, *DIO3* and *MEG3* were present on the microarray, and when their expression in AML samples was compared to that of SNORD114-3, the only strong correlation was seen in the case of *MEG3* (Fig. 1d in Study II). This positive correlation underlines the fact that only transcripts expressed from the same maternal allele of the locus are up-regulated. Furthermore, it implies that the expression of the whole maternally imprinted locus might be deregulated.

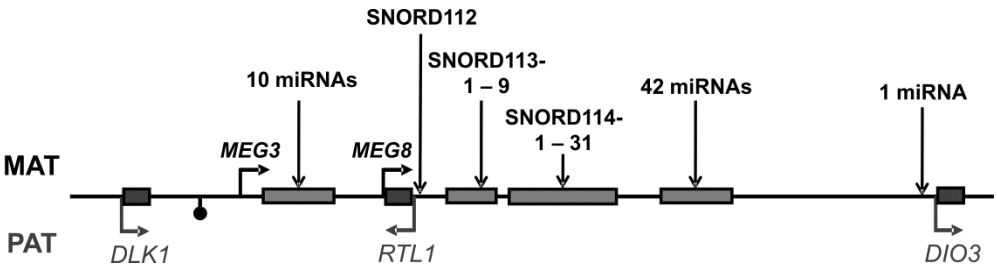

Figure 6. A schematic illustration of the imprinted *DLK1-DIO3* domain on 14q32.2. The transcripts expressed from the maternal allele (MAT) are given above the line representing the chromosome, and include the non-coding MEG3 and MEG8 (or RTL1-antisense), SNORD112, SNORD113 cluster (9 snoRNAs), SNORD114 cluster (31 snoRNAs), and in total 53 miRNAs. The paternal DLK-DIO3 allele (PAT) produces three protein-coding transcripts, DLK1, RTL1, and DIO3, shown below the line representing the chromosome. The black circle with a stem represents the intergenic imprinted control region of the domain. Illustration not drawn to scale. Data retrieved from www.ncbi.nlm.nih.gov/gene/.

5.2 The nucleolar protein composition in leukemia (III)

As previously described, the nucleolar protein composition is dynamic, and several nucleolar proteins have been implicated in cancer. In order to study the differences in the nucleolar proteomes between acute leukemia subtypes, we analyzed the protein contents of nucleolar extractions made from cell lines representing AML, T-ALL and pre-B-ALL. The analysis was carried out using two-dimensional difference gel electrophoresis (2-D DIGE), which is a method widely used to study complex protein mixtures. Its advantages include sensitivity, accurate quantification, comparability of numerous samples, and reliable detection of differences between samples, all due to the use of fluorescent labeling and an internal standard (Lilley and Friedman 2004).

Our strategy was to compare AML (OCI-AML3) with ALL (CCRF-CEM) as well as T-ALL (CCRF-CEM) with pre-B-ALL (MHH-CALL3). First, we prepared high-quality nucleolar extractions of the cell lines. In the 2-D DIGE experiments, each of the two samples was labeled with a separate fluorescent dye, combined

with the internal standard, and separated in the two dimensions. An example of the resulting gel is shown in Fig. 1 in Study III, in part A as fluorescent image and in part B as silver-stained gel, where the separated proteins can be seen as individual spots. The fluorescence intensity levels of the spots, which reflect the expression level of protein(s) present in the samples, were recorded.

Differences in expression between samples in each protein spot can be analyzed by comparing the intensity levels of the separate fluorescent dyes used for labeling the samples. The DeCyder software calculates the average ratio of the differentially expressed spots. When the cut-off for the average ratio was set at $>|2|$, 36 spots for AML/ALL comparison, and 29 for pre-B-/T-ALL comparison were detected. In order to identify the differentially expressed protein(s) present, the spots were excised from the gel and analyzed by mass spectrometry. Table 1 in Study III lists the proteins identified with high probability in the four spots of the pre-B-/T-ALL gel. To further confirm the results, two of the proteins, prohibitin (PHB) and TAR DNA-binding protein 43 (TDP-43) were studied with immunofluorescence and WB. Both PHB and TDP-43 were strongly expressed in the pre-B-ALL cell line MHH-CALL3 as compared to the T-ALL cell line CCRF-CEM (Fig. 2a-b and Fig. 3a-b in Study III), as also indicated by the 2-D DIGE analyses. In addition, both PHB and TDP-43 co-localized with the nucleolar marker NPM. Taken together, 2-D DIGE can be used to discern actual proteomic differences in the nucleoli of leukemia subtypes. To our knowledge, this was the first time when this method has been applied to analyzing leukemic nucleoli.

5.3 Functional studies in the zebrafish: SAP30L as an example (IV)

The identification of nucleolar proteins with altered expression between leukemia subtypes is certainly a start, but in order to illuminate their potential contribution to leukemogenesis, a step further has to be taken. While a wide range of *in vitro* experiments can be performed to shed light on protein function, they produce only small pieces of information at a time, which may turn out to be challenging to piece together. In order to gain an overall picture, the function of the protein of interest should be studied in the context of a living organism. Naturally, this requires sufficient background information regarding both the protein to be studied and the organism to be used. Also, a tool box for experimental analysis has

to be available. Our choice for the model organism was the zebrafish, particularly the embryos and larvae up to the age of seven days. A nuclear/nucleolar protein previously studied in our group was chosen to pioneer in setting up the system.

The Sin3A-associated protein 30-like (SAP30L) was first discovered in a screen carried out to identify novel mRNAs whose expression was induced by TGF-β (Lindfors et al. 2003). SAP30L shares 70 % sequence identity with another family member, SAP30, and both proteins are part of the Sin3A corepressor complex, a multiprotein cluster which performs the deacetylation of histones, and thus regulates gene expression (Silverstein and Ekwall 2005). The role of the SAP30 family proteins in the Sin3A complex is to link and stabilize the interactions between protein members and between proteins and DNA. Both SAP30L and SAP30 are basic and small, and have targeting signals to the nucleus and to the nucleolus whither they can also direct Sin3A. The major difference between these two proteins lies in the N-terminus, where SAP30 has an insertion of 38 amino acids relative to SAP30L (Viiri et al. 2006). Interestingly, the zebrafish is exceptional among organisms as it only has one member of the SAP30 family, the sequence of which more closely resembles that of human SAP30L than SAP30 (Viiri et al. 2009). It is thus the perfect model organism to study specifically SAP30L function *in vivo*, which has thus far remained unknown.

5.3.1 SAP30L expression in the zebrafish

To begin with, the expression of SAP30L was analyzed in zebrafish embryos and larvae, and in the tissues of the adult zebrafish. As shown in Fig. 1a-b in Study IV, SAP30L mRNA is expressed at various time-points during embryogenesis. The expression is localized in the brain at 24 hours post-fertilization (hpf) and onward, in the pectoral fin buds and later in the pectoral fins (from 36 hpf onward), in the common cardinal vein from 36 hpf onward, and in the pericardial cavity and the heart (from 24 hpf onward). Also, moderate expression is seen in the trunk, and in some embryos in the blood island at 24 hpf. In adult fish tissues, strong expression of SAP30L mRNA is seen in the brain, and moderate expression in muscle, eye, gonad and heart (see Fig. 1c in Study IV).

5.3.2 The effects of depleting SAP30L in the zebrafish

Next, the expression of SAP30L was knocked down during zebrafish early development using morpholino oligonucleotides (MO), which block the translation of SAP30L mRNA. Two independent SAP30L-targeting MOs were designed, and their efficacy was demonstrated *in vivo*, as SAP30L protein levels were reduced in MO-treated embryos (see Fig. 2 in Study IV). Also, SAP30L-MO1 inhibited the *in vitro* translation of SAP30L mRNA dose-dependently. SAP30L-MO1 or SAP30L-MO2 were injected into the embryos at 1 - 4 cell stage, and the resulting phenotype of the MO-injected zebrafish (morphants) was analyzed up to 5 - 6 days post-fertilization (dpf). In addition, a microarray analysis of gene expression in the morphants at 12 and 24 hpf was carried out to elucidate the transcriptional pathways SAP30L participates in. The phenotypes obtained with both SAP30L-MOs were essentially similar. Also, possible off-target effects mediated by p53, which are sometimes induced by MOs, were ruled out by co-injection with a p53-targeting MO.

5.3.2.1 SAP30L knock-down induces cardiac defects

The SAP30L morphants displayed considerable pericardial edema which progressed during the observation time. Moreover, almost 90 % of them showed deformed cardiac morphology compared to the control larvae (see Fig. 3a-b in Study IV). Cardiac function was also impaired: the morphants had significantly reduced heart rates, and the performance of the atrium and ventricle was substantially decreased as compared to the control larvae (Fig. 3c-d in Study IV). As SAP30L is expressed in the developing and adult zebrafish heart, the cardiac phenotype might well result from the knock-down of SAP30L expression. Indeed, one of the genes down-regulated at 12 hpf in SAP30L morphants, as revealed by the microarray analysis, was nkx2.5 (Fig. 4 in Study IV). nkx2.5 is known to regulate cardiac development (Chen and Fishman 1996, Targoff et al. 2008). The down-regulation of nkx2.5 in SAP30L-morphants was confirmed by RT-qPCR (Fig. 5a in Study IV). Furthermore, co-injection of nkx2.5 mRNA with SAP30L-MO1 partially rescued the deformed cardiac morphology. These observations suggest that in the zebrafish, nkx2.5 expression is regulated by a repressor complex containing SAP30L.

5.3.2.2 SAP30L knock-down evokes reduced hemoglobin levels

Some SAP30L morphants exhibited red blood cells paler than those in the control larvae. Subsequently, the hemoglobin levels were determined by o-dianisidine staining at 3 dpf, and the morphants showed reduced staining (Fig. 3e in Study IV), indicating reduced hemoglobin levels in the red blood cells. Consistent with this, microarray analysis of gene expression revealed a marked down-regulation of genes involved in hemoglobin synthesis and erythropoiesis, for example ppox and alas2 (Brownlie et al. 1998, Dailey and Dailey 1996), at 24 hpf (Fig. 4 in Study IV). The microarray result was further confirmed by RT-qPCR, which showed down-regulation of ppox expression in morphant embryos (Fig. 5c in Study IV).

5.3.2.3 Multiple signaling pathways are affected

The microarray analysis of gene expression in the SAP30L morphant and control larvae showed 205 genes to be up-regulated and 174 genes to be down-regulated at 12 hpf. At 24 hpf, 143 genes were up-regulated and 64 genes down-regulated (Fig. 4 in Study IV). Forty-six genes were identified at both time-points, and showed similar change in expression. Interestingly, the expression of several proteins involved in TGF-β signaling was affected by SAP30L knock-down, providing further support for the role of SAP30L in TGF-β signaling. It is noteworthy that several transcription factors were up- or down-regulated following SAP30L depletion. Taken together, SAP30L knock-down causes changes in genes involved in the regulation of transcription. The affected factors and pathways include TGF-β signaling, Wnt-family members, genes encoding mitochondrial proteins, as well as cardiogenesis, hemoglobin synthesis and erythropoiesis.

6. Discussion and future perspectives

6.1 Nucleolar components in leukemia

6.1.1 snoRNAs in leukemia

Relatively little is known regarding snoRNAs in leukemia at the general level. Recently, observations have been published linking snoRNAs and leukemia. For example, Valleron and associates reported the expression of snoRNAs to be globally down-regulated in ALL and AML compared to immuno-matched healthy cells, suggesting a different mode of regulation in normal and leukemic cells (Valleron et al. 2012a). On the other hand, reports associating individual snoRNAs with different types of cancer are available, as described in section 2.1.5.1. Taken together, it is noteworthy that both up- or down-regulation of expression of snoRNAs as well as mutations in snoRNAs have been reported in cancer.

In Study I, the whole snoRNA transcriptome of several leukemic cell lines was analyzed, and multiple snoRNAs with either up- or down-regulated expression between leukemia subtypes were discovered. Importantly, the differential expression of individual snoRNAs revealed by the SOLiD analysis could be validated by another sensitive method (RT-qPCR). However, apparent variation in snoRNA expression was seen within the leukemia subtypes. This probably reflects the specific characteristics of the individual cell lines, and is a typical drawback in using immortalized cell lines. In order to detect the real trends of differential snoRNA expression between leukemia subtypes, more cell lines as well as patient samples should be analyzed. As such, our observations provide a comprehensive systematic analysis of snoRNA expression in several subtypes of acute leukemia, and lend support to the conception that snoRNAs are differentially expressed in a leukemia subtype-specific manner. Of interest, it would appear that the choice of starting material may have an impact on results. As we observed, when total RNA was used instead of fractionated small RNA (under 200 nt), snoRNAs were only weakly detected.

Most of the nine snoRNAs studied have been described in the literature. scaRNA9 guides the 2´-O-ribose methylation of snRNA U2, and appears to be independently transcribed, as the transcript possesses the 5'-cap structure (Tycowski et al. 2004). SNORD24 has been discovered to be encoded in the intron of RPL7a (Qu et al. 1995), and SNORD55 in the intron of RPS8 (Nicoloso et al. 1996). SNORD82 resides in an intron of the nucleolin gene (Rebane and Metspalu 1999), and the discovery of SNORD105 (Vitali et al. 2003) and SNORD110 (Huttenhofer et al. 2001) have also been described. However, functional information is available only for SNORD44, which is encoded in the multi-snoRNA ncRNA host Gas5 (Smith and Steitz 1998), and has been shown to be associated with poor prognosis in cancer patients (Gee et al. 2011). Interestingly, this association was found when several snoRNAs (including SNORD44) often used as reference genes to normalize miRNA expression were evaluated. This, taken together with our observations on the differential expression of snoRNAs, underlines the fact that snoRNAs are not stably expressed and are therefore not suitable for use as reference genes. The ideal gene for normalizing expression data is constantly expressed in all samples included in the analysis.

Interestingly, in Study I, some bias towards scaRNAs among the up-regulated snoRNAs is to be seen. Similarly, box C/D appear to be over-represented among the down-regulated snoRNAs (see Table 10). Naturally it has to be taken into account that the overall number of DE snoRNAs identified is rather small, but there can be several additional reasons for these observations. One might be that box C/D snoRNAs are somewhat shorter, and might thus pass the 200 nt cut-off of the extraction kit more efficiently. Even if some box H/ACA snoRNAs are over 200 nt in size, they were firmly present in the SOLiD sequencing data, indicating that longer snoRNAs can also be detected. Furthermore, several scaRNAs were implicated to be differentially expressed. Most scaRNAs are over 200 nt in length, and thus their presence argues against the observed differences stemming from technical limitations. Presuming that our results on differential expression of snoRNAs between T-ALL and pre-B-ALL reflect the true situation in patients, some intriguing questions arise. Why are box H/ACA snoRNAs featured to a lesser degree? Do scaRNAs play a particular role in acute leukemias? Is the differential expression of snoRNAs related to the maturation status of the leukemic blast cells or the specific lineage? What are the mechanisms of action? Are the differences seen specific for acute leukemia subtypes, or are they equally

fundamental for other cancer types? Answering these and other questions will surely provide interesting information on the multiple roles of snoRNAs.

Also, in the comparison between T-ALL and pre-B-ALL, more up-regulated than down-regulated snoRNAs were seen. The same was seen when AML and ALL were compared. Since our analysis is based on comparisons made between different disease subtypes, it remains a matter of speculation whether the trend towards global down-regulation reported by Valleron et al. (Valleron et al. 2012a) would also be evident in the analyzed leukemic cells. As snoRNAs may operate as tumor-suppressors and oncogenes, it might be necessary to assess each snoRNA in a cellular context-dependent manner. In our limited set of samples, the snoRNA transcriptomes in leukemia subtypes differed from a solid cancer, represented by HeLa cells, but not extensively. This may suggest common mechanisms underlying the emergence of cancer, or, on the other way, reflect the maturation or differentiation status of the leukemic cells.

It is obvious, based on our selection of DE snoRNAs, that even though the genomic localization and possibly the modification target of many snoRNAs have been described, little is known as to their function. The conception that classical snoRNAs only guide modifications has predominated to such an extent that unprejudiced analyses aiming to reveal all potential functions of both classical and unorthodox snoRNAs have only recently started to emerge. Partly this has to be credited to exciting reports describing the roles of other ncRNAs in cancer, which have stimulated new ideas, partly again to the development of large-scale analysis techniques and coordinated sample libraries.

In Study II, we analyzed the expression of the 15 snoRNAs present on the microarray across pediatric acute leukemia samples and healthy controls. Contrary to the findings of Valleron et al. (Valleron et al. 2012a), these 15 snoRNAs did not show down-regulation of expression in leukemia as compared to normal samples. The reasons might lie in the number of snoRNAs analyzed (here: 15; Valleron et al.: almost all snoRNAs), or in the substantially larger number of patient samples included in our analysis. Similarly to us, Valleron et al. used a gene expression microarray in their preliminary analysis, suggesting that methodological differences may not be the key factor underlying the divergent observations. Nevertheless, similar results concerning the role of snoRNAs of the SNORD114 cluster in APL were also obtained in our analysis. Valleron et al. showed that many members of

this snoRNA cluster are up-regulated in APL compared to PML-RARA -negative AML, including SNORD114-3, which showed approximately three-fold higher expression (Valleron et al. 2012a). This is well in line with our observations. Furthermore, SNORD114-1 has proved to be involved in the modulation of cell growth and proliferation in APL cells (Valleron et al. 2012a). This suggests a role in leukemogenesis for snoRNAs, and in particular the SNORD114 cluster. As adumbrated here, one contributing factor could be the location of the SNORD114 cluster in the imprinted *DLK1-DIO3* domain, shown to be deregulated. Of note, it has been reported that the overexpression of the cluster is not dependent on the *PML-RARA* fusion gene (Cohen et al. 2012).

SNORD114-3, like all the members of the SNORD114 cluster, is an orphan snoRNA, without a known RNA target. The function of such snoRNAs is currently unknown, but regulatory roles have been suggested. These include functioning as precursors for miRNAs or other small RNA molecules (Ender et al. 2008, Ono et al. 2011, Scott et al. 2009, Scott and Ono 2011), or involvement in alternative mRNA splicing events (Bazeley et al. 2008, Scott et al. 2012). Another orphan snoRNA, ACA11, has been reported to be overexpressed in multiple myeloma patients, contributing to increased proliferation, suppression of oxidative stress, and chemoresistance (Chu et al. 2012). Discovering the function of SNORD114-3 is likely to be challenging as there are several highly similar snoRNAs present in the same cluster. To begin with, it would be interesting to analyze the cellular localization of this snoRNA as well as to identify the potential RNA, DNA and protein molecules that interact with it. This basic knowledge would already provide many hints on the function of SNORD114-3.

All in all, Study II provides an example of research where published data have been compiled and re-analyzed. As said, there is always only a limited number and amount of patient samples available, and therefore re-using patient-derived data is patient-sparing and reduces expenses, and enables the compilation of large datasets with numerous samples of even rare disease subtypes. One challenge is that although case histories are available, data may be inaccurately or inconsistently given, which complicates reliable categorization of samples, and also the wider analysis of other possibly contributing factors such as precise genotype, disease history, treatment regime used, age, family history, etc. Another challenge is that in order to be included, data have to be produced on identical platforms, and similar processing and analysis methods have to be employed. Even though the collection

and compilation of a database is time- and labor-consuming, and requires special expertise, many different hypotheses can ultimately be studied using these extensive patient data. In Study II, one drawback was the fact that only 15 snoRNAs could be included in the detailed analysis. On the other hand, the expression of these 15 snoRNAs was thoroughly analyzed in a notably large number of patient samples. Moreover, other genes of the *DLK1-DIO3* locus were present, and their expression status could be incorporated in the analysis. Thus, although it was not possible to conduct a comprehensive analysis of the expression of all snoRNAs in pediatric leukemias, interesting observations could be made when the phrasing of the research question was modified to match the limitations.

6.1.2 Application of snoRNAs in therapeutics and diagnostics

It has become evident that the ncRNA world holds numerous promises for application as diagnostic tools as well as therapeutic targets. miRNAs, which regulate gene expression post-transcriptionally, are probably the most extensively studied class of ncRNA in hematological malignancies (Lawrie 2013). miRNAs show different expression profiles in pediatric ALL, with various subgroups displaying distinct miRNA expression patterns, including both up- and down-regulated versions (de Oliveira et al. 2012). miRNAs can be stably detected from patient blood and saliva (Chen et al. 2008, Mitchell et al. 2008), which holds promise for their utilization as biomarkers (Hu et al. 2010, Ng et al. 2009, Tsujiura et al. 2010). However, translating these observations into actual diagnostic or therapeutic approaches has proved challenging. Both restoring lost miRNA expression (by means of miRNA mimics) and inhibiting overexpressed miRNAs are promising therapeutic strategies in leukemia (Murray et al. 2012).

Even though the mechanisms of action and molecular characteristics differ between miRNAs and snoRNAs, these two classes of ncRNA also have similarities. Importantly, both have been shown to be differentially expressed in leukemic cells when compared to their normal counterparts. Also, the division of the classes seems not to be clear-cut, as some snoRNAs are processed into miRNA-like molecules (Brameier et al. 2011, Ender et al. 2008), and some precursors of miRNAs resemble snoRNAs (Ono et al. 2011). Interestingly, snoRNAs can also be detected in the plasma (Liao et al. 2010). This suggests that similarly to miRNAs, also snoRNAs could be useful in the detection and

diagnostics of leukemia and other cancers. There are only a few reports describing the application of snoRNAs as biomarkers or therapeutic targets in leukemia. In peripheral T-cell lymphoma (PTCL), it has been shown that profiling snoRNA expression has diagnostic and prognostic significance, with subtypes of PTCL distinguishable by individual snoRNAs (Valleron et al. 2012b). Similarly, a set of snoRNAs has been identified to be down-regulated in CLL (Ronchetti et al. 2013). These snoRNAs could be useful as biomarkers to predict the potential outcome in patients at an early stage. It would appear that snoRNAs could be used in the diagnosis, prognosis and treatment of leukemia, similarly to miRNAs. As the function of snoRNAs is known to a lesser degree compared to miRNAs, the first applications could be in the prognosis or diagnosis of leukemia. If down-regulation of snoRNAs proves to be the general trend, as suggested (Valleron et al. 2012a), therapeutic approaches to re-induce or bring about ectopic expression of snoRNAs in order to restore the down-regulated expression have to be developed.

Next, it appears to be necessary to obtain more information on the snoRNA expression patterns in a leukemia subtype-specific manner. The analyses should be designed to answer crucial questions arising from the actual treatment of patients. In this way the results could be applied within a moderately short time. For example, the differences between relapsed and cured cases with regard to snoRNA expression could be studied. The effects on prognosis and disease progression, for example, will provide useful information. One challenge is to set up a standardized method for analysis which is both reliable and easily available. Further, efforts should be made to gain a deeper understanding of the many aspects of snoRNA function, as this would provide clues as to how to target the deregulated or dysfunctional snoRNAs in order to restore their normal function. Currently, it seems that fragments of information are available on many issues or phenomena, but no single phenomenon is known completely to the core. This leaves much room for imprecision and guessing, which is not desirable when working with patients. Nevertheless, application of the observations "from bench to bedside" has to start somewhere, even with limited information. One could envision the first clinical applications for snoRNAs in leukemia to be in the identification of special cases among specific subtypes which e.g. are more prone to relapse, or respond to treatment differently. At the moment it seems that the interesting journey has just begun.

6.1.3 Nucleolar proteins in leukemia

As already noted, several nucleolar proteins with different functions have been implicated in cancer. Since many proteins reside in the nucleolus only occasionally and shuttle in and out of it, we were interested in finding differences in nucleolar protein composition in various acute leukemia subtypes. According to our reasoning, this might reveal novel players involved in acute leukemia pathogenesis in a subtype-specific manner, or provide potential biomarkers. We chose two-dimensional difference gel electrophoresis as the method, since this allows analysis of the whole nucleolar proteome of a sample at once, and differentially expressed proteins can thus be identified. As our results demonstrate, marked alterations in nucleolar protein compositions can be detected using this method. Another advantage of this approach is that comparison of samples from different runs is possible due to the internal standard. In the case of acute leukemia, with multiple specific disease subtypes, this method could be deployed to define both general and subtype-specific alterations in the nucleolar protein composition.

The two identified proteins, PHB and TDP-43, have many reported functions in the cell, and are also implicated in diseases. PHB, initially described to be a mitochondrial chaperone, is shown to co-localize with Rb and p53 in the nucleus, and to modulate their transcriptional activity (Fusaro et al. 2003, Wang et al. 2002). Originally PHB was considered to be a tumor suppressor (Fusaro et al. 2003, Wang et al. 1999). However, several observations attribute PHB both anti- and pro-tumorigenic functions, the decisive factor being the sub-cellular localization (Theiss and Sitaraman 2011). The nucleolar roles for PHB could be similar to those in the nucleus and mitochondria, for example regulation of transcription or chaperoning ribosome biogenesis. On the other hand, PHB could be detained in the nucleolus, which leads to inhibition of its functions in other cellular locations. TDP-43, a nuclear ribonucleoprotein, is involved in pre-mRNA splicing and regulation of transcription (Chen-Plotkin et al. 2010), has an established role in neurodegenerative diseases(Arai et al. 2006, Neumann et al. 2006), and is involved in the regulation of the cell cycle and increased apoptosis via Rb (Ayala et al. 2008). Moreover, TDP-43 plays a role in the post-transcriptional regulation of miRNA expression (Kawahara and Mieda-Sato 2012). In principle, TDP-43 could participate in leukemogenesis via many of the reported roles, since deregulation of Rb or miRNAs has been reported in cancer, for example.

Taken together, based on previous reported observations, both proteins could contribute in many ways to the emergence of acute leukemia. Interestingly, both PHB and TDP-43 are involved in transcriptional repression, PHB via involvement with transcriptional repression complexes (Wang et al. 2002), and TDP-43 presumably by binding the promoter region of specific genes (Buratti and Baralle 2010). Disturbed regulation of transcription is typical of many cancers. Also, the change in cellular localization contributing to disease seems to be a common mechanism for both proteins. It should be noted that strong expression of PHB and TDP-43 in the nucleolus is not unique to pre-B-ALL, as the other tested leukemic cell lines also expressed PHB and TDP-43 strongly (See Fig. 2a-b and Fig. 3a-b in Study III). In fact, the only cell line with weak expression of PHB and TDP-43 was CCRF-CEM, which was used in the 2-D DIGE analysis. This suggests that, again, the cell line-specific characteristics influence the results, and more samples should be analyzed in order to decipher the possible relevance of these proteins in acute leukemia subtypes. Also, the possible contribution of other factors has to be taken into account.

One important limitation in analyzing nucleolar proteome is the large amount of starting material required for the extraction, which may hinder the analysis of patient samples. In addition, the 2-D DIGE is a time- and labor-consuming technique involving a great deal of handwork. There are also limitations common to all two-dimensional gel electrophoresis (2-DE) analyses combined with MS, for example neglecting proteins with very high or low molecular weight, and insensitivity to detect low-abundance proteins. Furthermore, a common feature emerging from the 2-DE analyses carried out to identify differentially expressed proteins is that some proteins and protein families are repeatedly implicated regardless of the experimental set-up or species. In point of fact, the protein families of both PHB and TDP-43 are listed among those recurrently identified in comparative proteomics. This phenomenon is not fully understood, but probably technical factors, relative protein abundance, and the possible function of the protein in response to cellular conditions all contribute (Petrak et al. 2008, Wang et al. 2009). Nevertheless, while keeping these limitations in mind, and when the results are proportioned with other related observations, proteomic analyses of the nucleoli in specific leukemia subtypes will provide novel information in abundance. For example, proteomic profiling of the T-cell nucleolus has provided lineage-specific insights into the nucleolar protein composition (Jarboui et al. 2011).

6.2 The zebrafish as a model for nucleolar protein function

Use of the zebrafish as a model to shed light on the *in vivo* function of the nucleolar/nuclear protein SAP30L yielded some interesting observations. As we employed zebrafish embryos and larvae in our study, we were able to follow the effects of SAP30L depletion on embryogenesis and early development. Knock-down of gene expression by MOs is not a perfect technique, as off-target effects have been reported, for example, but it is widely used also in other organisms, especially in developmental biology. In any case, when properly controlled, results are reliable. Initially we focused on the observed phenotype, which strikingly turned out to involve cardiac deformation and malfunction as well as defects in the hemoglobinization of red blood cells. Analysis of the morphants was feasible, as the larvae are transparent, and for example the heart could be videoed and its function measured in the living organism. Also, many analysis methods have been published for the zebrafish, which, at least in our case, could be set up with relatively small effort and cost. In order to look deeper into the effects of SAP30L depletion, we carried out a microarray analysis of the gene expression, which revealed many affected pathways. The fact that in our case only SAP30L, and not the related SAP30, was present in the zebrafish was a clear advantage. The effects of permanent SAP30L knock-down would be interesting to follow through to adulthood. Keeping in mind that many of the larvae were very unfit at 6-7 dpf, it is probable that here not many would have survived into adulthood. In addition, as the techniques for permanent gene silencing operate at the genetic level, the resulting phenotype might have been slightly different.

The zebrafish offers an excellent opportunity to study the previously unknown function of a protein in the context of a living organism. When planning a knock-down experiment, extra care should be taken in designing the morpholinos, and also in ascertaining the effect *in vivo*. In the case of a nucleolar protein, the effects of depleting the protein will be seen at the level of the whole animal, tissues and all cellular compartments, as it is not possible to target the knock-down specifically to the nucleolus. For example in the case of SAP30L, it is not possible to designate the morphant phenotype on one specific function or localization of the protein. It is, however, possible to analyze the cells at the nucleolar level. Furthermore, the morphant phenotype might provide clues to the nucleolar effect. If there is no hypothesis as to what the morphant phenotype could be like, it might be

challenging to perceive it in its entirety. Reporter fish lines, with a certain cell type fluorescing, might provide assistance in this. On the other hand, if the protein under study is relatively well-known, the zebrafish also offers a platform to study specific questions. Of note, an effort to produce inactive mutant fish lines for all the protein-coding genes in the zebrafish genome is under way (Kettleborough et al. 2013). This initiative, with far-reaching insight, will surely expedite and contribute to multiple interesting discoveries. One of the eminent advantages of the zebrafish is the possibility it affords to analyze cellular events in a living organism in real time and in a versatile manner.

7. Conclusions

This study was carried out in order to shed light on the role of nucleolar components in acute leukemia. The major conclusions are as follows:

1. snoRNAs are differentially expressed between cell lines representing major acute leukemia subtypes.

2. a hematological database combining data produced with a common platform provides an excellent means to study the expression of snoRNAs in a large cohort of patients.

3. the overexpression of SNORD114-3 is seen in APL.

4. 2-D DIGE is an applicable proteomic method to identify differentially expressed nucleolar proteins between acute leukemia subtypes.

5. the zebrafish provides a versatile *in vivo* model to study the function of a nucleolar protein.

6. SAP30L is involved in the regulation of cardiogenesis and hematopoiesis during zebrafish early development.

8. Acknowledgments

This study was carried out at the Tampere Center for Child Health Research (formerly the Paediatric Research Centre), School of Medicine, University of Tampere and Tampere University Hospital, Finland during 2009-2013. I wish to thank Professor Per Ashorn, Professor Matti Korppi and the Chief of the Department of Pediatrics Merja Helminen for providing excellent working facilities. I also want to acknowledge the contribution of the Institute of Biomedical Technology (IBT), University of Tampere, in collaboration regarding equipment and professional expertise, with special thanks to the Zebrafish Core Facility at IBT. Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB) is acknowledged for organizing interesting courses and seminars, and for a travel grant. I am very thankful for the financial support from Onerva Sinisalo legacy funds. This series was financially supported by the Competitive Research Funding of Tampere University Hospital, University of Tampere, the Foundation of Paediatric Research in Finland, the Academy of Finland Research Council for Health, the Nona and Kullervo Väre Foundation, and the Päivikki and Sakari Sohlberg Foundation.

I am truly grateful to my supervisors Docent Olli Lohi, MD, PhD, and Professor Markku Mäki for the opportunity to work with interesting research topics in an inspiring work community. Olli, it is impossible to put into words how much your enthusiasm, insight and positive attitude have influenced me as a scientist and also as a person. During these years I have learned much from you. I find it especially important and motivating that the research questions we work on stem from actual clinical problems. You also have the remarkable ability to simplify matters or at least make them appear simple, regarding both science and practical issues. All this makes working with you inconceivably easy and enjoyable. Markku, I wish to thank you for your support and guidance during my thesis project. Please, keep asking (the difficult) questions! I have been fortunate with my divergent research projects which have challenged me and introduced me to numerous techniques and people. Markku and Olli, you have taught me the importance and power of collaboration. Thank you both for everything.

The official reviewers of my thesis, Docent Maija Vihinen-Ranta, PhD, and Docent Matti Korhonen, MD, PhD, are acknowledged for their constructive critical comments and suggestions which improved this thesis. Members of my thesis committee, Docent Mataleena Parikka, DDS, PhD, and Docent Marko Pesu, MD, PhD, are warmly thanked for their guidance and advice regarding my projects. I am happy that I also got to work with both of you.

I want to express my warmest gratitude to all the coauthors and collaborators for their contribution regarding the original publications of this thesis. It would have been impossible to carry out the research without your contribution. I would especially like to thank Asta Laiho, MSc, and Thomas Liuksiala, BSc, for their invaluable contribution and for patiently answering my numerous questions regarding mysterious bioinformatics processes. Taisto Heinonen, PhD, is acknowledged for improving the language of the original publications, and Robert McGilleon, MA, is thanked for revising the language of this thesis. A warm thank you goes to Anu Jantunen, MSc, for revising the Finnish abstract.

Past and present members of the Hemato-Oncology Research Group HemoRes and Markku's Coeliac Disease Study Group are thanked for help and pleasant coexistence in the lab. I warmly thank my office roommates and fellow HemoRessians Toni Grönroos, MSc, and Susanna Teppo, MSc, for nice discussions, giving opinions, and tolerating me during the writing of this thesis. Laura Airaksinen, MSc, Cristina Nadalutti, PhD, Suvi Kalliokoski, MSc, Tiina Rauhavirta, MSc, Hanna Korkeamäki, MSc, MB, Heidi Kontro, MSc, Ana-Marija Sulic, PhD, Minna Hietikko, MSc, and Yashavanthi Niranjan, MSc, are thanked for peer support and the fun times off and on duty. Mr. Jorma Kulmala and Annemari Uusimäki, BHSc, are warmly acknowledged for all their help and for sharing cheerful moments when working together. FM-3 3rd floor folk are jointly acknowledged for help in practical matters and nice (coffee room) conversations. Having been around at FM-1/2/3 since 2002, there are a lot of people who have helped me over the years. So, a collective thank you goes to all the fellow students, professors, supervisors, colleagues, technicians, lab neighbors and corridor acquaintances☺.

I wish to thank all my dear friends for the great and unforgettable times we have had together. Having life outside the lab is important and it is all due to you. Thank you for sharing your lives with me. You are all precious to me. The

common denominator of some of my friends is Riihimäki, which means that I am fortunate in having known you for many (many) years. I especially want to thank Petra and Antti for the fun trips in Finland and abroad, the mind-blowing culinary pleasures and always being ready for action. Thank you Sanna for asking me to sit next to you on the first day in school. Thank you Outi for giving me different perspectives on life. I also wish to acknowledge the balancing factors in my life without which it would have been very difficult to carry out this work: literature and cinema (providing escape from everyday (lab) life), BodyCombat (a way to release frustration and aggression), and inspirational music. If my thesis had a soundtrack, it would include many tracks by Bruce Springsteen and Tori Amos.

I would like to express my sincere gratitude to my family for their support. Kiitos äiti ja isä että olette uskoneet minuun, kannustaneet eteenpäin elämässä ja auttaneet kun olen apua tarvinnut. Kiitos että olen saanut tehdä omat päätökseni. Minä en olisi minä ilman teitä. Kiitos myös isoäideilleni Meerille ja Tertulle kaikesta mitä olen vuosien varrella teiltä oppinut ja saanut. Meeri-mummolle erityiskiitokset leipomiskipinän sytyttämisestä. My dear cousin Mari has been in charge of my hair and jewelry for many years. Thank you for keeping me in touch with my roots. To me, you are the embodiment of perseverance. I also wish to thank my family-in-law for their interest in my work. Kiitokset appivanhemmilleni Vuokolle ja Reinolle kaikista mukavista hetkistä ja ihanista tekstiileistä.

My darling husband Markus, thank you for your unconditional love and support. Everything would lack significance without you. I love you.

Tampere, January 5th, 2014

Kaisa Teittinen

9. References

Abdelmohsen K and Gorospe M. (2012) RNA-binding protein nucleolin in disease. RNA Biol. 9, 799-808.

Ahmad Y, Boisvert FM, Gregor P, Cobley A and Lamond AI. (2009) NOPdb: Nucleolar Proteome Database--2008 update. Nucleic Acids Res. 37, D181-4.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K and Walter P. (2002) Molecular Biology of the Cell. (New York: Garland Science).

Anders S and Huber W. (2010) Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI and Mann M. (2005) Nucleolar proteome dynamics. Nature 433, 77-83.

Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, Mann M and Lamond AI. (2002) Directed proteomic analysis of the human nucleolus. Curr. Biol. 12, 1-11.

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y and Oda T. (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem. Biophys. Res. Commun. 351, 602-611.

Askarian-Amiri ME, Crawford J, French JD, Smart CE, Smith MA, Clark MB, Ru K, Mercer TR, Thompson ER, Lakhani SR, Vargas AC, Campbell IG, Brown MA, Dinger ME and Mattick JS. (2011) SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. RNA 17, 878-891.

Audas TE, Jacob MD and Lee S. (2012a) Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. Mol. Cell 45, 147-157.

Audas TE, Jacob MD and Lee S. (2012b) The nucleolar detention pathway: A cellular strategy for regulating molecular networks. Cell. Cycle 11, 2059-2062.

Ayala YM, Misteli T and Baralle FE. (2008) TDP-43 regulates retinoblastoma protein phosphorylation through the repression of cyclin-dependent kinase 6 expression. Proc. Natl. Acad. Sci. U. S. A. 105, 3785-3789.

Bartova E, Horakova AH, Uhlirova R, Raska I, Galiova G, Orlova D and Kozubek S. (2010) Structure and epigenetics of nucleoli in comparison with non-nucleolar compartments. J. Histochem. Cytochem. 58, 391-403.

Bateman CM, Colman SM, Chaplin T, Young BD, Eden TO, Bhakta M, Gratias EJ, van Wering ER, Cazzaniga G, Harrison CJ, Hain R, Ancliff P, Ford AM, Kearney L and Greaves M. (2010) Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. Blood 115, 3553-3558.

Bazeley PS, Shepelev V, Talebizadeh Z, Butler MG, Fedorova L, Filatov V and Fedorov A. (2008) snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. Gene 408, 172-179.

Belin S, Beghin A, Solano-Gonzalez E, Bezin L, Brunet-Manquat S, Textoris J, Prats AC, Mertani HC, Dumontet C and Diaz JJ. (2009) Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. PLoS One 4, e7147.

Bhojwani D and Pui CH. (2013) Relapsed childhood acute lymphoblastic leukaemia. Lancet Oncol. 14, e205-17.

Boisvert FM, van Koningsbruggen S, Navascues J and Lamond AI. (2007) The multifunctional nucleolus. Nat. Rev. Mol. Cell Biol. 8, 574-585.

Boulon S, Westman BJ, Hutten S, Boisvert FM and Lamond AI. (2010) The nucleolus under stress. Mol. Cell 40, 216-227.

Boyd MT, Vlatkovic N and Rubbi CP. (2011) The nucleolus directly regulates p53 export and degradation. J. Cell Biol. 194, 689-703.

Brameier M, Herwig A, Reinhardt R, Walter L and Gruber J. (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic Acids Res. 39, 675-686.

Brownlie A, Donovan A, Pratt SJ, Paw BH, Oates AC, Brugnara C, Witkowska HE, Sassa S and Zon LI. (1998) Positional cloning of the zebrafish sauternes gene: a model for congenital sideroblastic anaemia. Nat. Genet. 20, 244-250.

Bueno C, Montes R, Catalina P, Rodriguez R and Menendez P. (2011) Insights into the cellular origin and etiology of the infant pro-B acute lymphoblastic leukemia with MLL-AF4 rearrangement. Leukemia 25, 400-410.

Buratti E and Baralle FE. (2010) The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. RNA Biol. 7, 420-429.

Burger K and Eick D. (2013) Functional ribosome biogenesis is a prerequisite for p53 destabilization: impact of chemotherapy on nucleolar functions and RNA metabolism. Biol. Chem. 394, 1133-1143.

Cancer Genome Atlas Research Network. (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. 368, 2059-2074.

Cassidy SB, Schwartz S, Miller JL and Driscoll DJ. (2012) Prader-Willi syndrome. Genet. Med. 14, 10-26.

Chen JN and Fishman MC. (1996) Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. Development 122, 3809-3816.

Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, Zhang Y, Wang J, Jiang X, Xiang Y, Xu C, Zheng P, Zhang J, Li R, Zhang H, Shang X, Gong T, Ning G, Wang J, Zen K, Zhang J and Zhang CY. (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res. 18, 997-1006.

Chen-Plotkin AS, Lee VM and Trojanowski JQ. (2010) TAR DNA-binding protein 43 in neurodegenerative disease. Nat. Rev. Neurol. 6, 211-220.

Chu L, Su MY, Maggi LB,Jr, Lu L, Mullins C, Crosby S, Huang G, Chng WJ, Vij R and Tomasson MH. (2012) Multiple myeloma-associated chromosomal translocation activates orphan snoRNA ACA11 to suppress oxidative stress. J. Clin. Invest. 122, 2793-2806.

Cohen Y, Hertzog K, Reish O, Mashevich M, Garach-Jehoshua O, Bar-Chaim A, Trakhtenbrot L and Kornberg A. (2012) The increased expression of 14q32 small nucleolar RNA transcripts in promyelocytic leukemia cells is not dependent on PML-RARA fusion gene. Blood Cancer. J. 2, e92.

Collins K. (2006) The biogenesis and regulation of telomerase holoenzymes. Nat. Rev. Mol. Cell Biol. 7, 484-494.

Colombo E, Alcalay M and Pelicci PG. (2011) Nucleophosmin and its complex network: a possible therapeutic target in hematological diseases. Oncogene 30, 2595-2609.

Coute Y, Burgess JA, Diaz JJ, Chichester C, Lisacek F, Greco A and Sanchez JC. (2006) Deciphering the human nucleolar proteome. Mass Spectrom. Rev. 25, 215-234.

Dailey TA and Dailey HA. (1996) Human protoporphyrinogen oxidase: expression, purification, and characterization of the cloned enzyme. Protein Sci. 5, 98-105.

Darzacq X, Jady BE, Verheggen C, Kiss AM, Bertrand E and Kiss T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. EMBO J. 21, 2746-2756.

De Keersmaecker K, Atak ZK, Li N, Vicente C, Patchett S, Girardi T, Gianfelici V, Geerdens E, Clappier E, Porcu M, Lahortiga I, Luca R, Yan J, Hulselmans G, Vranckx H, Vandepoel R, Sweron B, Jacobs K, Mentens N, Wlodarska I, Cauwelier B, Cloos J, Soulier J, Uyttebroeck A, Bagni C, Hassan BA, Vandenberghe P, Johnson AW, Aerts S and Cools J. (2013) Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. Nat. Genet. 45, 186-190.

de Oliveira JC, Brassesco MS, Scrideli CA, Tone LG and Narendran A. (2012) MicroRNA expression and activity in pediatric acute lymphoblastic leukemia (ALL). Pediatr. Blood Cancer. 59, 599-604.

de The H, Le Bras M and Lallemand-Breitenbach V. (2012) The cell biology of disease: Acute promyelocytic leukemia, arsenic, and PML bodies. J. Cell Biol. 198, 11-21.

Decatur WA and Fournier MJ. (2002) rRNA modifications and ribosome function. Trends Biochem. Sci. 27, 344-351.

Decatur WA and Fournier MJ. (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. J. Biol. Chem. 278, 695-698.

Derenzini M, Montanaro L and Trere D. (2009) What the nucleolus says to a tumour pathologist. Histopathology 54, 753-762.

Derenzini M, Trere D, Pession A, Govoni M, Sirri V and Chieco P. (2000) Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. J. Pathol. 191, 181-186.

Derenzini M, Trere D, Pession A, Montanaro L, Sirri V and Ochs RL. (1998) Nucleolar function and size in cancer cells. Am. J. Pathol. 152, 1291-1297.

Di Bacco A, Ouyang J, Lee HY, Catic A, Ploegh H and Gill G. (2006) The SUMO-specific protease SENP5 is required for cell division. Mol. Cell. Biol. 26, 4489-4498.

Dieci G, Preti M and Montanini B. (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. Genomics 94, 83-88.

Dong XY, Guo P, Boyd J, Sun X, Li Q, Zhou W and Dong JT. (2009) Implication of snoRNA U50 in human breast cancer. J. Genet. Genomics 36, 447-454.

Dong XY, Rodriguez C, Guo P, Sun X, Talbot JT, Zhou W, Petros J, Li Q, Vessella RL, Kibel AS, Stevens VL, Calle EE and Dong JT. (2008) SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. Hum. Mol. Genet. 17, 1031-1042.

Drexler HG, Matsuo AY and MacLeod RA. (2000) Continuous hematopoietic cell lines as model systems for leukemia-lymphoma research. Leuk. Res. 24, 881-911.

ENCODE Project Consortium, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C and Snyder M. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57-74.

Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N and Meister G. (2008) A human snoRNA with microRNA-like functions. Mol. Cell 32, 519-528.

Estey E and Dohner H. (2006) Acute myeloid leukaemia. Lancet 368, 1894-1907.

Ferreira HJ, Heyn H, Moutinho C and Esteller M. (2012) CpG island hypermethylation-associated silencing of small nucleolar RNAs in human cancer. RNA Biol. 9, 881-890.

Fusaro G, Dasgupta P, Rastogi S, Joshi B and Chellappan S. (2003) Prohibitin induces the transcriptional activity of p53 and is exported from the nucleus upon apoptotic signaling. J. Biol. Chem. 278, 47853-47861.

Gee HE, Buffa FM, Camps C, Ramachandran A, Leek R, Taylor M, Patil M, Sheldon H, Betts G, Homer J, West C, Ragoussis J and Harris AL. (2011) The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. Br. J. Cancer 104, 1168-1177.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY and Zhang J. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 5, R80.

Gerbi SA, Borovjagin AV and Lange TS. (2003) The nucleolus: a site of ribonucleoprotein maturation. Curr. Opin. Cell Biol. 15, 318-325.

Girardot M, Cavaille J and Feil R. (2012) Small regulatory RNAs controlled by genomic imprinting and their contribution to human disease. Epigenetics 7, 1341-1348.

Goodell MA and Godley LA. (2013) Perspectives and future directions for epigenetics in hematology. Blood 121, 5131-5137.

Greaves M and Maley CC. (2012) Clonal evolution in cancer. Nature 481, 306-313.

Greco A. (2009) Involvement of the nucleolus in replication of human viruses. Rev. Med. Virol. 19, 201-214.

Grisendi S, Mecucci C, Falini B and Pandolfi PP. (2006) Nucleophosmin and cancer. Nat. Rev. Cancer. 6, 493-505.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, Kelsh RN, Furutani-Seiki M, Vogelsang E, Beuchle D, Schach U, Fabian C and Nusslein-Volhard C. (1996) The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. Development 123, 1-36.

Hanahan D and Weinberg RA. (2011) Hallmarks of cancer: the next generation. Cell 144, 646-674.

Hein N, Hannan KM, George AJ, Sanij E and Hannan RD. (2013) The nucleolus: an emerging target for cancer therapy. Trends Mol. Med.

Heinäniemi M, Nykter M, Kramer R, Wienecke-Baldacchino A, Sinkkonen L, Zhou JX, Kreisberg R, Kauffman SA, Huang S and Shmulevich I. (2013) Gene-pair expression signatures reveal lineage control. Nat. Methods 10, 577-583.

Hernandez-Verdun D. (2011) Assembly and disassembly of the nucleolus during the cell cycle. Nucleus 2, 189-194.

Hiscox JA, Whitehouse A and Matthews DA. (2010) Nucleolar proteomics and viral infection. Proteomics 10, 4077-4086.

Hu Z, Chen X, Zhao Y, Tian T, Jin G, Shu Y, Chen Y, Xu L, Zen K, Zhang C and Shen H. (2010) Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. J. Clin. Oncol. 28, 1721-1726.

Huttenhofer A, Kiefmann M, Meier-Ewert S, O'Brien J, Lehrach H, Bachellerie JP and Brosius J. (2001) RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. EMBO J. 20, 2943-2953.

Inaba H, Greaves M and Mullighan CG. (2013) Acute lymphoblastic leukaemia. Lancet 381, 1943-1955.

Jack K, Bellodi C, Landry DM, Niederer RO, Meskauskas A, Musalgaonkar S, Kopmar N, Krasnykh O, Dean AM, Thompson SR, Ruggero D and Dinman JD. (2011) rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. Mol. Cell 44, 660-666.

Jan M and Majeti R. (2013) Clonal evolution of acute leukemia genomes. Oncogene 32, 135-140.

Jarboui MA, Wynne K, Elia G, Hall WW and Gautier VW. (2011) Proteomic profiling of the human T-cell nucleolus. Mol. Immunol. 49, 441-452.

Kawahara Y and Mieda-Sato A. (2012) TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. Proc. Natl. Acad. Sci. U. S. A. 109, 3347-3352.

Kennedy JA and Barabe F. (2008) Investigating human leukemogenesis: from cell lines to in vivo models of human leukemia. Leukemia 22, 2029-2040.

Kettleborough RN, Busch-Nentwich EM, Harvey SA, Dooley CM, de Bruijn E, van Eeden F, Sealy I, White RJ, Herd C, Nijman IJ, Fenyes F, Mehroke S, Scahill C, Gibbons R, Wali N, Carruthers S, Hall A, Yen J, Cuppen E and Stemple DL. (2013) A systematic genome-wide analysis of zebrafish protein-coding gene function. Nature 496, 494-497.

Kiss T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell 109, 145-148.

Kiss T, Fayet E, Jady BE, Richard P and Weber M. (2006) Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. Cold Spring Harb. Symp. Quant. Biol. 71, 407-417.

Kiss T, Fayet-Lebaron E and Jady BE. (2010) Box H/ACA small ribonucleoproteins. Mol. Cell 37, 597-606.

Kohler A and Hurt E. (2007) Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol. 8, 761-73.

Korkeamäki H, Viiri K, Kukkonen MK, Maki M and Lohi O. (2008) Alternative mRNA splicing of SAP30L regulates its transcriptional repression activity. FEBS Lett. 582, 379-384.

Kurki S, Peltonen K, Latonen L, Kiviharju TM, Ojala PM, Meek D and Laiho M. (2004) Nucleolar protein NPM interacts with HDM2 and protects tumor suppressor protein p53 from HDM2-mediated degradation. Cancer. Cell. 5, 465-475.

Lafontaine DL and Tollervey D. (2001) The function and synthesis of ribosomes. Nat. Rev. Mol. Cell Biol. 2, 514-520.

Lawrie CH. (2013) MicroRNAs in hematological malignancies. Blood Rev. 27, 143-154.

Lee JT and Gu W. (2010) The multiple levels of regulation by p53 ubiquitination. Cell Death Differ. 17, 86-92.

Lestrade L and Weber MJ. (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic Acids Res. 34, D158-62.

Li Z and Hann SR. (2013) Nucleophosmin is essential for c-Myc nucleolar localization and c-Myc-mediated rDNA transcription. Oncogene 32, 1988-1994.

Liang XH, Liu Q and Fournier MJ. (2009) Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. RNA 15, 1716-1728.

Liao J, Yu L, Mei Y, Guarnera M, Shen J, Li R, Liu Z and Jiang F. (2010) Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. Mol. Cancer. 9, 198.

Lilley KS and Friedman DB. (2004) All about DIGE: quantification technology for differential-display 2D-gel proteomics. Expert Rev. Proteomics 1, 401-409.

Lindfors K, Viiri KM, Niittynen M, Heinonen TY, Maki M and Kainulainen H. (2003) TGF-beta induces the expression of SAP30L, a novel nuclear protein. BMC Genomics 4, 53.

Liu B, Liang XH, Piekna-Przybylska D, Liu Q and Fournier MJ. (2008) Mis-targeted methylation in rRNA can severely impair ribosome synthesis and activity. RNA Biol. 5, 249-254.

Liu JM and Ellis SR. (2006) Ribosomes and marrow failure: coincidental association or molecular paradigm? Blood 107, 4583-4588.

Lohi O, Parikka M and Ramet M. (2013) The zebrafish as a model for paediatric diseases. Acta Paediatr. 102, 104-110.

Lowenberg B, Downing JR and Burnett A. (1999) Acute myeloid leukemia. N. Engl. J. Med. 341, 1051-1062.

Lymberopoulos MH, Bourget A, Ben Abdeljelil N and Pearson A. (2011) Involvement of the UL24 protein in herpes simplex virus 1-induced dispersal of B23 and in nuclear egress. Virology 412, 341-348.

MacLeod RA, Nagel S, Scherr M, Schneider B, Dirks WG, Uphoff CC, Quentmeier H and Drexler HG. (2008) Human leukemia and lymphoma cell lines as models and resources. Curr. Med. Chem. 15, 339-359.

Maggi LB,Jr and Weber JD. (2005) Nucleolar adaptation in human cancer. Cancer Invest. 23, 599-608.

Mannoor K, Liao J and Jiang F. (2012) Small nucleolar RNAs in cancer. Biochim. Biophys. Acta 1826, 121-128.

Marcel V, Ghayad SE, Belin S, Therizols G, Morel AP, Solano-Gonzalez E, Vendrell JA, Hacot S, Mertani HC, Albaret MA, Bourdon JC, Jordan L, Thompson A, Tafer Y, Cong R, Bouvet P, Saurin JC, Catez F, Prats AC, Puisieux A and Diaz JJ. (2013) p53 Acts as a Safeguard of Translational Control by Regulating Fibrillarin and rRNA Methylation in Cancer. Cancer. Cell. 24, 318-330.

Masetti R, Biagi C, Zama D, Vendemini F, Martoni A, Morello W, Gasperini P and Pession A. (2012) Retinoids in pediatric onco-hematology: the model of acute promyelocytic leukemia and neuroblastoma. Adv. Ther. 29, 747-762.

McStay B and Grummt I. (2008) The epigenetics of rRNA genes: from molecular to chromosome biology. Annu. Rev. Cell Dev. Biol. 24, 131-157.

Mei YP, Liao JP, Shen JP, Yu L, Liu BL, Liu L, Li RY, Ji L, Dorsey SG, Jiang ZR, Katz RL, Wang JY and Jiang F. (2011) Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. Oncogene

Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB and Tewari M. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proc. Natl. Acad. Sci. U. S. A. 105, 10513-10518.

Mongelard F and Bouvet P. (2007) Nucleolin: a multiFACeTed protein. Trends Cell Biol. 17, 80-86.

Montanaro L, Trere D and Derenzini M. (2008) Nucleolus, ribosomes, and cancer. Am. J. Pathol. 173, 301-310.

Montanaro L, Trere D and Derenzini M. (2012) Changes in ribosome biogenesis may induce cancer by down-regulating the cell tumor suppressor potential. Biochim. Biophys. Acta 1825, 101-110.

Moore HM, Bai B, Boisvert FM, Latonen L, Rantanen V, Simpson JC, Pepperkok R, Lamond AI and Laiho M. (2011) Quantitative proteomics and dynamic imaging of the nucleolus reveal distinct responses to UV and ionizing radiation. Mol. Cell. Proteomics 10, M111.009241.

Mori H, Colman SM, Xiao Z, Ford AM, Healy LE, Donaldson C, Hows JM, Navarrete C and Greaves M. (2002) Chromosome translocations and covert leukemic clones are generated during normal fetal development. Proc. Natl. Acad. Sci. U. S. A. 99, 8242-8247.

Mourtada-Maarabouni M, Pickard MR, Hedge VL, Farzaneh F and Williams GT. (2009) GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene 28, 195-208.

Mullighan CG. (2012) Molecular genetics of B-precursor acute lymphoblastic leukemia. J. Clin. Invest. 122, 3407-3415.

Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, Shurtleff SA and Downing JR. (2008) Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science 322, 1377-1380.

Murray MY, Rushworth SA and MacEwan DJ. (2012) Micro RNAs as a new therapeutic target towards leukaemia signalling. Cell. Signal. 24, 363-368.

Nallar SC, Lin L, Srivastava V, Gade P, Hofmann ER, Ahmed H, Reddy SP and Kalvakolanu DV. (2011) GRIM-1, a Novel Growth Suppressor, Inhibits rRNA Maturation by Suppressing Small Nucleolar RNAs. PLoS One 6, e24082.

Narla A and Ebert BL. (2010) Ribosomopathies: human disorders of ribosome dysfunction. Blood 115, 3196-3205.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ and Lee VM. (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314, 130-133.

Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, Poon TC, Ng SS and Sung JJ. (2009) Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. Gut 58, 1375-1381.

Nicoloso M, Qu LH, Michot B and Bachellerie JP. (1996) Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2'-O-ribose methylation of rRNAs. J. Mol. Biol. 260, 178-195.

Ono M, Scott MS, Yamada K, Avolio F, Barton GJ and Lamond AI. (2011) Identification of human miRNA precursors that resemble box C/D snoRNAs. Nucleic Acids Res. 39, 3879-3891.

Payne E and Look T. (2009) Zebrafish modelling of leukaemias. Br. J. Haematol. 146, 247-256.

Petrak J, Ivanek R, Toman O, Cmejla R, Cmejlova J, Vyoral D, Zivny J and Vulpe CD. (2008) Deja vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins. Proteomics 8, 1744-1749.

Pui CH. (2006) Childhood leukemias. (United Kingdom: Cambridge University Press).

Pui CH. (2010) Recent research advances in childhood acute lymphoblastic leukemia. J. Formos. Med. Assoc. 109, 777-787.

Pui CH, Carroll WL, Meshinchi S and Arceci RJ. (2011a) Biology, risk stratification, and therapy of pediatric acute leukemias: an update. J. Clin. Oncol. 29, 551-565.

Pui CH, Gajjar AJ, Kane JR, Qaddoumi IA and Pappo AS. (2011b) Challenging issues in pediatric oncology. Nat. Rev. Clin. Oncol. 8, 540-549.

Pui CH, Mullighan CG, Evans WE and Relling MV. (2012) Pediatric acute lymphoblastic leukemia: where are we going and how do we get there? Blood 120, 1165-1174.

Pui CH, Robison LL and Look AT. (2008) Acute lymphoblastic leukaemia. Lancet 371, 1030-1043.

Qu LH, Henry Y, Nicoloso M, Michot B, Azum MC, Renalier MH, Caizergues-Ferrer M and Bachellerie JP. (1995) U24, a novel intron-encoded small nucleolar RNA with two 12 nt long, phylogenetically conserved complementarities to 28S rRNA. Nucleic Acids Res. 23, 2669-2676.

R Development Core Team. (2008). R: A Language and Environment for Statistical Computing.(Vienna, Austria: R Foundation for Statistical Computing).

Radtke I, Mullighan CG, Ishii M, Su X, Cheng J, Ma J, Ganti R, Cai Z, Goorha S, Pounds SB, Cao X, Obert C, Armstrong J, Zhang J, Song G, Ribeiro RC, Rubnitz JE, Raimondi SC, Shurtleff SA and Downing JR. (2009) Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia. Proc. Natl. Acad. Sci. U. S. A. 106, 12944-12949.

Ransom DG, Haffter P, Odenthal J, Brownlie A, Vogelsang E, Kelsh RN, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Mullins MC and Nusslein-Volhard C. (1996) Characterization of zebrafish mutants with defects in embryonic hematopoiesis. Development 123, 311-319.

Rebane A and Metspalu A. (1999) U82, a novel snoRNA identified from the fifth intron of human and mouse nucleolin gene. Biochim. Biophys. Acta 1446, 426-430.

Reichow SL, Hamma T, Ferre-D'Amare AR and Varani G. (2007) The structure and function of small nucleolar ribonucleoproteins. Nucleic Acids Res. 35, 1452-1464.

Richard P and Kiss T. (2006) Integrating snoRNP assembly with mRNA biogenesis. EMBO Rep. 7, 590-592.

Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA and Ekker SC. (2007) P53 Activation by Knockdown Technologies. PLoS Genet. 3, e78.

Ronchetti D, Mosca L, Cutrona G, Tuana G, Gentile M, Fabris S, Agnelli L, Ciceri G, Matis S, Massucco C, Colombo M, Reverberi D, Recchia AG, Bossio S, Negrini M, Tassone P, Morabito F, Ferrarini M and Neri A. (2013) Small nucleolar RNAs as new biomarkers in chronic lymphocytic leukemia. BMC Med. Genomics 6, 27.

Rubnitz JE. (2012) How I treat pediatric acute myeloid leukemia. Blood 119, 5980-5988.

Ruggero D and Pandolfi PP. (2003) Does the ribosome translate cancer? Nat. Rev. Cancer. 3, 179-192.

Sasaki M, Kawahara K, Nishio M, Mimori K, Kogo R, Hamada K, Itoh B, Wang J, Komatsu Y, Yang YR, Hikasa H, Horie Y, Yamashita T, Kamijo T, Zhang Y, Zhu Y, Prives C, Nakano T, Mak TW, Sasaki T, Maehama T, Mori M and Suzuki A. (2011) Regulation of the MDM2-P53 pathway and tumor growth by PICT1 via nucleolar RPL11. Nat. Med. 17, 944-951.

Scherl A, Coute Y, Deon C, Calle A, Kindbeiter K, Sanchez JC, Greco A, Hochstrasser D and Diaz JJ. (2002) Functional proteomic analysis of human nucleolus. Mol. Biol. Cell 13, 4100-4109.

Schmittgen TD, Jiang J, Liu Q and Yang L. (2004) A high-throughput method to monitor the expression of microRNA precursors. Nucleic Acids Res. 32, e43.

Schmittgen TD and Livak KJ. (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. 3, 1101-1108.

Schneider DA. (2012) RNA polymerase I activity is regulated at multiple steps in the transcription cycle: recent insights into factors that influence transcription elongation. Gene 493, 176-184.

Scott MS, Avolio F, Ono M, Lamond AI and Barton GJ. (2009) Human miRNA precursors with box H/ACA snoRNA features. PLoS Comput. Biol. 5, e1000507.

Scott MS, Boisvert FM, McDowall MD, Lamond AI and Barton GJ. (2010) Characterization and prediction of protein nucleolar localization sequences. Nucleic Acids Res. 38, 7388-7399.

Scott MS and Ono M. (2011) From snoRNA to miRNA: Dual function regulatory non-coding RNAs. Biochimie 93, 1987-1992.

Scott MS, Ono M, Yamada K, Endo A, Barton GJ and Lamond AI. (2012) Human box C/D snoRNA processing conservation across multiple cell types. Nucleic Acids Res. 40, 3676-3688.

Sherr CJ. (2006) Divorcing ARF and p53: an unsettled case. Nat. Rev. Cancer. 6, 663-673.

Sherr CJ and Weber JD. (2000) The ARF/p53 pathway. Curr. Opin. Genet. Dev. 10, 94-99.

Silverstein RA and Ekwall K. (2005) Sin3: a flexible regulator of global gene expression and genome stability. Curr. Genet. 47, 1-17.

Sirri V, Urcuqui-Inchima S, Roussel P and Hernandez-Verdun D. (2008) Nucleolus: the fascinating nuclear body. Histochem. Cell Biol. 129, 13-31.

Smith CM and Steitz JA. (1998) Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. Mol. Cell. Biol. 18, 6897-6909.

Smyth GK. (2005) Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions using R and Bioconductor, R. Gentleman, V. Carey, S. Dudoit, and W. Huber R. Irizarry ed., (New York: Springer) pp. 397-398-420.

Stoletov K and Klemke R. (2008) Catch of the day: zebrafish as a human cancer model. Oncogene 27, 4509-4520.

Su H, Xu T, Ganapathy S, Shadfan M, Long M, Huang TH, Thompson I and Yuan ZM. (2013) Elevated snoRNA biogenesis is essential in breast cancer. Oncogene doi: 10.1038/onc.2013.89

Sundqvist A, Liu G, Mirsaliotis A and Xirodimas DP. (2009) Regulation of nucleolar signalling to p53 through NEDDylation of L11. EMBO Rep. 10, 1132-1139.

Szczepanski T, Harrison CJ and van Dongen JJ. (2010) Genetic aberrations in paediatric acute leukaemias and implications for management of patients. Lancet Oncol. 11, 880-889.

Tafforeau L, Zorbas C, Langhendries JL, Mullineux ST, Stamatopoulou V, Mullier R, Wacheul L and Lafontaine DL. (2013) The Complexity of Human Ribosome Biogenesis Revealed by Systematic Nucleolar Screening of Pre-rRNA Processing Factors. Mol. Cell 51, 539-551.

Tang R, Dodd A, Lai D, McNabb WC and Love DR. (2007) Validation of zebrafish (Danio rerio) reference genes for quantitative real-time RT-PCR normalization. Acta Biochim. Biophys. Sin. (Shanghai) 39, 384-390.

Targoff KL, Schell T and Yelon D. (2008) Nkx genes regulate heart tube extension and exert differential effects on ventricular and atrial cell number. Dev. Biol. 322, 314-321.

Teittinen KJ, Gronroos T, Parikka M, Ramet M and Lohi O. (2012) The zebrafish as a tool in leukemia research. Leuk. Res. 36, 1082-1088.

Theiss AL and Sitaraman SV. (2011) The role and therapeutic potential of prohibitin in disease. Biochim. Biophys. Acta

Thisse C and Thisse B. (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat. Protoc. 3, 59-69.

Toledo F and Wahl GM. (2006) Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. Nat. Rev. Cancer. 6, 909-923.

Trinkle-Mulcahy L, Chusainow J, Lam YW, Swift S and Lamond A. (2007) Visualization of intracellular PP1 targeting through transiently and stably expressed fluorescent protein fusions. Methods Mol. Biol. 365, 133-154.

Trinkle-Mulcahy L and Lamond AI. (2006) Mitotic phosphatases: no longer silent partners. Curr. Opin. Cell Biol. 18, 623-631.

Tsujiura M, Ichikawa D, Komatsu S, Shiozaki A, Takeshita H, Kosuga T, Konishi H, Morimura R, Deguchi K, Fujiwara H, Okamoto K and Otsuji E. (2010) Circulating microRNAs in plasma of patients with gastric cancers. Br. J. Cancer 102, 1174-1179.

Tycowski KT, Aab A and Steitz JA. (2004) Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa. Curr. Biol. 14, 1985-1995.

Vagnarelli P, Hudson DF, Ribeiro SA, Trinkle-Mulcahy L, Spence JM, Lai F, Farr CJ, Lamond AI and Earnshaw WC. (2006) Condensin and Repo-Man-PP1 co-operate in the regulation of chromosome architecture during mitosis. Nat. Cell Biol. 8, 1133-1142.

Valleron W, Laprevotte E, Gautier EF, Quelen C, Demur C, Delabesse E, Agirre X, Prosper F, Kiss T and Brousset P. (2012a) Specific small nucleolar RNA expression profiles in acute leukemia. Leukemia. 26, 2052–2060.

Valleron W, Ysebaert L, Berquet L, Fataccioli V, Quelen C, Martin A, Parrens M, Lamant L, de Leval L, Gisselbrecht C, Gaulard P and Brousset P. (2012b) Small nucleolar RNA expression profiling identifies potential prognostic markers in peripheral T-cell lymphoma. Blood. 120, 3997-4005.

van der Linden MH, Creemers S and Pieters R. (2012) Diagnosis and management of neonatal leukaemia. Semin. Fetal. Neonatal Med. 17, 192-195.

Van Vlierberghe P and Ferrando A. (2012) The molecular basis of T cell acute lymphoblastic leukemia. J. Clin. Invest. 122, 3398-3406.

Viiri KM, Heinonen TY, Maki M and Lohi O. (2009) Phylogenetic analysis of the SAP30 family of transcriptional regulators reveals functional divergence in the domain that binds the nuclear matrix. BMC Evol. Biol. 9, 149.

Viiri KM, Korkeamaki H, Kukkonen MK, Nieminen LK, Lindfors K, Peterson P, Maki M, Kainulainen H and Lohi O. (2006) SAP30L interacts with members of the Sin3A corepressor complex and targets Sin3A to the nucleolus. Nucleic Acids Res. 34, 3288-3298.

Vitali P, Royo H, Seitz H, Bachellerie JP, Huttenhofer A and Cavaille J. (2003) Identification of 13 novel human modification guide RNAs. Nucleic Acids Res. 31, 6543-6551.

Vlatkovic N, Boyd MT and Rubbi CP. (2013) Nucleolar control of p53: a cellular Achilles' heel and a target for cancer therapy. Cell Mol. Life Sci.

Wang P, Bouwman FG and Mariman EC. (2009) Generally detected proteins in comparative proteomics--a matter of cellular stress response? Proteomics 9, 2955-2966.

Wang S, Fusaro G, Padmanabhan J and Chellappan SP. (2002) Prohibitin co-localizes with Rb in the nucleus and recruits N-CoR and HDAC1 for transcriptional repression. Oncogene 21, 8388-8396.

Wang S, Nath N, Adlam M and Chellappan S. (1999) Prohibitin, a potential tumor suppressor, interacts with RB and regulates E2F function. Oncogene 18, 3501-3510.

Warner JR and McIntosh KB. (2009) How common are extraribosomal functions of ribosomal proteins? Mol. Cell 34, 3-11.

Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, Wartman LD, Lamprecht TL, Liu F, Xia J, Kandoth C, Fulton RS, McLellan MD, Dooling DJ, Wallis JW, Chen K, Harris CC, Schmidt HK, Kalicki-Veizer JM, Lu C, Zhang Q, Lin L, O'Laughlin MD, McMichael JF, Delehaunty KD, Fulton LA, Magrini VJ, McGrath SD, Demeter RT, Vickery TL, Hundal J, Cook LL, Swift GW, Reed JP, Alldredge PA, Wylie TN, Walker JR, Watson MA, Heath SE, Shannon WD, Varghese N, Nagarajan R, Payton JE, Baty JD, Kulkarni S, Klco JM, Tomasson MH, Westervelt P, Walter MJ, Graubert TA, DiPersio JF, Ding L, Mardis ER and Wilson RK. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150, 264-278.

Westerfield, M. (1995). The zebrafish book: A guide to the laboratory use of Zebrafish (Eugene, OR.: University of Oregon Press).

Williams GT and Farzaneh F. (2012) Are snoRNAs and snoRNA host genes new players in cancer? Nat. Rev. Cancer. 12, 84-88.

Williams GT, Mourtada-Maarabouni M and Farzaneh F. (2011) A critical role for non-coding RNA GAS5 in growth arrest and rapamycin inhibition in human T-lymphocytes. Biochem. Soc. Trans. 39, 482-486.

Wong JM, Kusdra L and Collins K. (2002) Subnuclear shuttling of human telomerase ind0uced by transformation and DNA damage. Nat. Cell Biol. 4, 731-736.

Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH and Dunn MJ. (2000) A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis 21, 3666-3672.

Zhai W and Comai L. (2000) Repression of RNA polymerase I transcription by the tumor suppressor p53. Mol. Cell. Biol. 20, 5930-5938.

Zhang Y and Lu H. (2009) Signaling to p53: ribosomal proteins find their way. Cancer. Cell. 16, 369-377.

10. Original communications

Cell Oncol. (2013) 36:55–63
DOI 10.1007/s13402-012-0113-5

ORIGINAL PAPER

Expression of small nucleolar RNAs in leukemic cells

Kaisa J. Teittinen · Asta Laiho · Annemari Uusimäki · Juha-Pekka Pursiheimo · Attila Gyenesei · Olli Lohi

Accepted: 5 November 2012 / Published online: 15 November 2012

Abstract
Purpose Small nucleolar RNAs (snoRNAs) direct sequence-specific modifications to ribosomal RNA. We hypothesized that the expression of snoRNAs may be altered in leukemic cells.
Methods The expression of snoRNAs was analyzed in various leukemic cell lines by massive parallel sequencing (SOLiD). Quantitative real-time PCR (RT-qPCR) was used to validate the expression profiles.
Results Our results show characteristic differences in the expression patterns of snoRNAs between cell lines representing the main subgroups of leukemia, AML, pre-B-ALL and T-ALL, respectively. In RT-qPCR analyses, several snoRNAs were found to be differentially expressed in T-ALL as compared to pre-B-ALL cell lines.
Conclusions snoRNAs are differentially expressed in various leukemic cell lines and could, therefore, be potentially useful in the classification of leukemia subgroups.

Keywords Leukemia · snoRNA · SOLiD · Nucleolus

Abbreviations

ALL	Acute lymphoblastic leukemia
AML	Acute myeloid leukemia
APL	Acute promyeloid leukemia
BLL	Burkitt's lymphoma/leukemia
DE	Differentially expressed
FC	Fold change
miRNA	Micro-RNA
ncRNA	Non-protein-coding RNA
NSCLC	Non-small cell lung cancer
RT-qPCR	Quantitative real-time PCR
scaRNA	Cajal body-specific RNA
siRNA	Small interfering RNA
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein

1 Introduction

The discovery that the majority of the mammalian genome is transcribed into non-protein-coding RNAs (ncRNAs) has changed the landscape of human genetics [1]. Recent work has established that these RNAs are involved in normal cellular processes as well as pathological conditions such as cancer, heart diseases, immune disorders, and various neurodegenerative and metabolic diseases [2]. Several classes of ncRNA have been identified, including small interfering RNAs (siRNA), microRNAs (miRNA), PIWI-associated RNAs (piRNA), small nucleolar RNAs (snoRNA), and various species of long ncRNA.

The small non-coding RNA molecules of the nucleolus, snoRNAs, can be grouped into two main classes. One class

Electronic supplementary material The online version of this article (doi:10.1007/s13402-012-0113-5) contains supplementary material, which is available to authorized users.

K. J. Teittinen (✉) · A. Uusimäki · O. Lohi
Tampere Center for Child Health Research, University of Tampere School of Medicine and Tampere University Hospital, FM-3, Biokatu 10, 3rd Floor,
FIN-33520 Tampere, Finland
e-mail: kaisa.teittinen@uta.fi

A. Laiho · J.-P. Pursiheimo · A. Gyenesei
Turku Centre for Biotechnology and the Finnish Microarray and Sequencing Centre, University of Turku and Åbo Akademi University, 20520 Turku, Finland

comprises the C/D box snoRNAs, which share the evolutionarily conserved C (RUGAUGA) and D (uCUGA) box motifs. The function of these RNAs is to guide methylation of specific target RNAs, and they associate with the nucleolar protein fibrillarin, which catalyzes the transfer of 2′-O-methyl groups [3]. Members of the other class, the H/ACA box snoRNAs, share the conserved box H (AnAnnA) and ACA motifs. These RNAs guide pseudouridylation by associating with dyskerin, the enzyme responsible for catalyzing the uridine-pseudouridine isomerization reaction [3]. An interesting group of snoRNAs, the small Cajal body-specific RNAs (scaRNAs), contains composite RNA molecules that guide both methylation and pseudouridylation of the spliceosomal small nuclear RNAs (snRNAs) [4, 5].

snoRNAs associate with a number of protein components to form the functional small nucleolar ribonucleoprotein (snoRNP) complex, which then carries out the RNA modification reactions. The snoRNAs are responsible for the selection of target sequences by transient complementary base-pairing with the target RNA molecules, which are in most cases ribosomal RNAs (rRNAs). In addition to directing the modification of rRNA, snoRNPs modify other RNAs, including spliceosomal snRNAs, some transfer RNAs (tRNA), and messenger RNAs (mRNA) [6].

In vertebrates, the majority of both C/D and H/ACA box snoRNAs are encoded within introns of protein-coding genes [7–10], whose protein products in many cases have some role in protein synthesis [11]. A common feature of the genes that host intronic snoRNAs is that most of them encode 5′ terminal oligopyrimidine (5′TOP) mRNAs, which are classically thought to produce proteins associated with translation [12, 13]. Intronic encoding provides an economical mechanism for simultaneous production of both the protein and RNA components needed in protein synthesis. After splicing of the pre-mRNA, the snoRNAs are removed from the host introns by exonucleases and processed further [10, 14–16]. While most snoRNAs have been found in protein-coding host genes, some reside in genes with no apparent protein-coding potential [12, 17, 18]. Such genes are presumed to act as vehicles for snoRNA production, although a recent report described additional snoRNA-independent functions for the snoRNA-host RNA, *Zfas1* [19]. Indeed, the expression strategies of snoRNAs show great diversity, ranging from different types of independent coding units with specialized promoters to intronic coding units without independent promoters. Encoding of snoRNAs can be organized in clusters or individually, and frequently a single host gene may encode several snoRNAs in adjacent introns, according to the "one snoRNA per intron" principle [20]. Thus, the expression strategies and genomic organization of snoRNAs are variable and complex, and our understanding of them is far from complete.

The general features of the archetypal snoRNAs have been known for some time. However, in addition to the stereotypic snoRNAs, recent work has identified several snoRNAs with structural and/or functional features typically associated with other types of ncRNA, adding a further layer of complexity. For example, mature snoRNAs can be further processed to produce short RNAs with miRNA-like functions [21–23], suggesting a connection between processing and silencing of RNAs. As new studies shed light on the relationships between different ncRNAs, unexpected complexities and unorthodox features are gradually emerging in the interplay between these RNA molecules.

The Next-generation sequencing technologies have played a crucial role in the high-throughput detection of ncRNAs. In the present work, we analyzed the expression patterns of snoRNAs in several leukemic cell lines using SOLiD, a massive parallel sequencing platform. Our results show that the expression patterns of snoRNAs comply with the classical subtyping of leukemias based on morphology and flow cytometry. The expression levels of individual snoRNAs analyzed by RT-qPCR confirm those revealed by SOLiD sequencing. Our results suggest that snoRNA profiling bears potential as a diagnostic tool for leukemia. Revealing the mechanisms that regulate snoRNA expression, and their possible role in pathogenesis, is of considerable interest and warrants further studies.

2 Materials and methods

2.1 Cell culture

Several cell lines were used. The cell lines representing pre-B-ALL were MHH-CALL3, KOPN-8, Kasumi-2, Nalm-6 and REH. The cell lines representing T-ALL were CCRF-CEM, Molt-16, Jurkat, P12-Ichikawa and Peer. The OCI-AML3 and HL-60 cell lines represented AML. The Daudi cell line represented Burkitt's lymphoma, and the HeLa cell line represented cervical carcinoma. All cell lines were obtained from the Leibniz-Institut DSMZ (German Collection of Microorganisms and Cell Culture, Braunschweig, Germany). KOPN-8, Kasumi-2, Nalm-6, REH, CCRF-CEM, Molt-16, Jurkat, P12-Ichikawa, HL-60 and Daudi cells were cultured in RPMI-1640 medium (Gibco, Invitrogen) supplemented with 1 % L-glutamine (L-glu), 10 % fetal bovine serum (FBS), and 1 % penicillin/streptomycin antibiotics (pen-strep). MHH-CALL3 and Peer cells were cultured in the same medium containing 20 % FBS. OCI-AML3 cells were cultured in MEM Alpha Medium (Gibco, Invitrogen) supplemented with 1 % L-glu, 20 % FBS and 1 % pen-strep. HeLa cells were cultured in D-MEM medium (Sigma-Aldrich) supplemented with 1 % L-glu, non-essential amino acids, 10 % FBS and 1 % pen-strep.

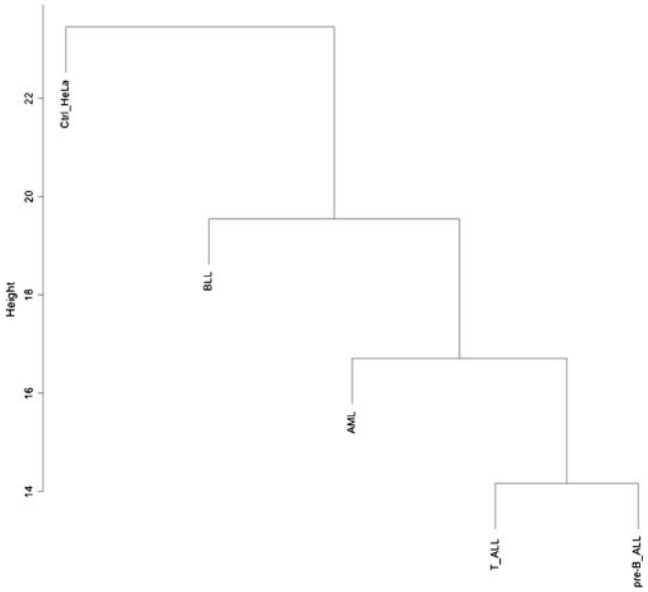

Fig. 1 Hierarchical clustering with Euclidean metrics for normalized snoRNA sequencing data. The tag count values are averaged for each subgroup of acute leukemia. Sample groups that share most similarity can be found in the branches nearest each other

2.2 Isolation of RNA

The Small RNA fractions (less than 200 nucleotides in length), which contain the snoRNAs, were isolated using the *mir*Vana miRNA isolation kit (Ambion, Applied Biosystems) according to the manufacturer's instructions. A method that selectively isolates smaller-sized RNA molecules was deployed in order to produce material highly enriched in snoRNAs for SOLiD sequencing analysis, enabling the detection of even low-expressing snoRNAs. As some snoRNAs are larger than 200 nt, they might be excluded and, thus, potentially bias the analysis. However, several snoRNAs larger than 200 nt in size were detected in our data set. For the isolation of RNA, cells were grown to 80–90 % confluence and counted, after which 10 million cells were collected by centrifugation and washed once in PBS. The quality and concentration of the RNA preparations were checked using a NanoDrop spectrophotometer (Thermo Fisher Scientific). Total RNA was extracted from each cell line using the *mir*Vana miRNA isolation kit (Ambion, Applied Biosystems) as instructed by the manufacturer.

2.3 Preparing libraries for SOLiD sequencing

The libraries for SOLiD sequencing were constructed following the SOLiD small RNA expression kit protocol provided by the manufacturer. The snoRNA sample was first hybridized and ligated in the Adaptor Mix. The RNA population containing the ligated adaptors was reverse-transcribed, purified with the MinElute PCR Purification Kit (Qiagen) and run on a denaturing gel, after which the cDNA of the desired size was excised. The size-selected cDNA was amplified with in-gel PCR, and cleaned up using the PureLink PCR Micro Kit (Invitrogen). The libraries were quantitated using a Qubit fluorimeter (Invitrogen) and quantitative PCR. The template bead preparation was performed by emulsion PCR (ePCR), and after enrichment of the template beads were deposited onto a sequencing glass slide.

2.4 SOLiD sequencing

The samples were processed on a SOLiD 4 instrument (Life Technologies) as instructed by the manufacturer. The read

Table 1 Characteristics of the cell lines used

cell line	type	patient origin	karyotype, fusion genes	analysis
KASUMI-2	pre-B ALL	15 years, diagnosis	hypodiploid	RT-qPCR
KOPN-8	pre-B ALL	3 months	hypodiploid, MLL-ENL	SOLiD, RT-qPCR
MHH-CALL3	pre-B ALL	11 years, diagnosis	pseudodiploid	SOLiD, RT-qPCR
NALM-6	pre-B ALL	19 years, relapse	near diploid	RT-qPCR
REH	pre-B ALL	15 years, first relapse	pseudodiploid, TEL-AML1	RT-qPCR
CCRF-CEM	T-cell ALL	3 years, terminal relapse	near tetraploid	SOLiD, RT-qPCR
JURKAT	T-cell ALL	14 years, first relapse	hypotetraploid	RT-qPCR
MOLT-16	T-cell ALL	5 years, relapse	near diploid	SOLiD, RT-qPCR
P12-Ichikawa	T-cell ALL	7 years	hypotetraploid	RT-qPCR
PEER	T-cell ALL	4 years, second relapse	pseudodiploid, NUP214-ABL1	RT-qPCR
HL-60	AML-M2	35 years	hypotetraploid/hypodiploid	SOLiD
OCI-AML3	AML-M4	57 years, diagnosis	hyperdiploid, NPM mutation	SOLiD
DAUDI	Burkitt lymphoma	16 years, diagnosis	near diploid	SOLiD
HeLa	cervix carcinoma	31 years	hypertriploid/hypotetraploid	SOLiD

ALL acute lymphoblastic leukemia; *AML* acute myeloid leukemia; *RT-qPCR* quantitative real-time PCR

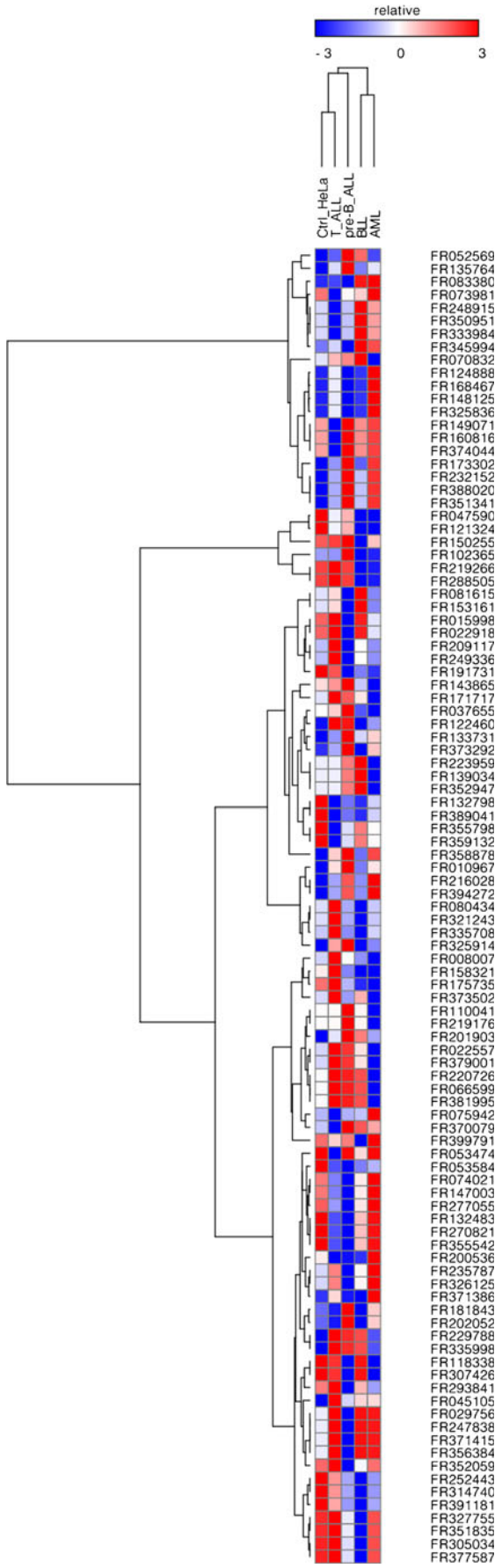

Fig. 2 A heat map clustering of the differentially expressed features for the comparisons between subgroups of leukemia. Pearson's metrics was used in the hierarchical clustering of the samples and filtered features. The clustering is based on the general expression measurement similarity. In the plot, red color denotes high expression and blue low expression. Each row represents one differentially expressed feature, and each column represents one sample subgroup

length was 35 bp. Fragment libraries were prepared for each sample. For SOLiD sequencing, two independent RNA samples were analyzed from each cell line.

2.5 Analysis of data

The data obtained were mapped to the human reference genome (hg18) using the SOLiD BioScope v.1.3 software package. The features of snoRNA against which the tags were counted, were retrieved from the fRNAdb, a non-coding RNA sequence database (http://www.ncrna.org/frnadb/) v. 3.4, which contained 981 snoRNA sequences. The downstream analysis was carried out using the R/Bioconductor software and the tag count data were normalized using the DESeq analysis package [24]. The data were found to be of high quality and the Pearson's correlations between samples in different sample groups were between 0.8–0.96, which is indicative of a high reproducibility. Statistical testing between groups was carried out using the Limma package, and features with a p-value below 0.05 and absolute fold change larger than 1.5 were considered as differentially expressed. The GENE-E tool (http://www.broadinstitute.org/cancer/software/GENE-E/) was used for the heatmap clustering of the data.

2.6 Quantitative real-time PCR

The expression of individual snoRNAs was measured using quantitative real-time PCR (RT-qPCR). The small RNA fraction and total RNA were isolated using the *mir*Vana miRNA isolation kit (Ambion, Applied Biosystems), and converted to cDNA using iScript Select cDNA Synthesis Kit (Bio-Rad) and random primers according to the manufacturers' instructions. The pooled templates for snoRNA analyses were prepared by mixing equal amounts of cDNA from each cell line in a given group. First we analyzed the cell lines that had been used in the SOLiD sequencing. The pooled T-ALL template consisted of cDNA derived from CCRF-CEM and Molt-16 cells, and the pooled pre-B-ALL template included cDNA from KOPN-8 and MHH-CALL3. Then, all pre-B and T leukemic cell lines were analyzed individually. U6 RNA was used for normalization of the data [25, 26]. The primer sequences are described in the Supplementary Material. RT-qPCR was performed by using the Evagreen Ssofast supermix kit (Bio-Rad) and a CFX96 real-time thermal cycler (Bio-Rad) as instructed by the manufacturer. An average threshold cycle (Ct) value was

calculated from two replicate samples. The runs were repeated twice or thrice, and no significant differences between replicate runs were observed. When appropriate, differences between groups were compared by using the two-tailed Student *t* test. A p-value of less than 0.05 was considered statistically significant.

3 Results and discussion

In order to study the expression of snoRNAs in leukemic cells, fractions representing small RNAs from various leukemic cell lines were analyzed by massive parallel sequencing (SOLiD 4, Life Technologies), (see Supplementary Fig. 1). The chosen cell lines included representatives of acute myeloid leukemia (AML), acute lymphoblastic leukemias (T-ALL and pre-B-ALL), Burkitt's lymphoma/leukemia (BLL), and a solid tumor (HeLa cell line) (see Table 1). The data were analyzed to detect global differences in the expression values of snoRNAs. In our dataset, hierarchical clustering of samples based on tag count values grouped the cell lines according to the leukemia subgroup (Fig. 1). These results show that different cancers and subgroups of leukemia have unique expression patterns of snoRNAs, and that various subgroups of leukemia are more closely related to each other than to a solid tumor (cervical carcinoma represented by HeLa cells).

In order to find differences among various leukemia subgroups and not individual cell lines, we compared the differentially expressed (DE) features between the subgroups (i.e. AML, pre-B-ALL, T-ALL) by using the fold change (FC) values and p-values. For filtering out differentially expressed genes, the minimum fold change was set at 1.5, and the p-value significance level at 0.05. This yielded 46 differentially expressed snoRNAs in the comparison of T-ALL with pre-B-ALL, of which 29 were up-regulated and 17 down-regulated. C/D box snoRNAs were affected the most: 19 of the up-regulated and 15 of the down-regulated snoRNAs were of the C/D box type. Interestingly, there were several scaRNAs among the up-regulated snoRNAs, the most prominent of which was scaRNA6. In the comparison of ALL (including T-ALL and pre-B-ALL cells) with AML, 40 differentially expressed snoRNAs were found, of which 22 were up-regulated and 18 down-regulated. In this comparison, the C/D box snoRNAs also stood out: 17 of the up-regulated and all of the down-regulated were C/D box snoRNAs. A heat map clustering of the DE features is shown in Fig. 2. These results indicate that the expression pattern of snoRNAs is dependent on the subgroup of leukemia, and that various subgroups of leukemia show marked differences in their snoRNA expression patterns.

Next, we focused on the differentially expressed snoRNAs from the T-ALL/pre-B-ALL comparison. In order to validate the expression profiles obtained by SOLiD analysis, the expression of nine individual snoRNAs was analyzed by RT-qPCR. These nine snoRNAs included both up- and down-regulated snoRNAs, and two scaRNAs (Table 2). Independent small RNA fractions were prepared from the T- and pre-B-ALL cell lines. First, we analyzed the leukemic cell lines used in the original SOLiD analysis. The RT-qPCR analyses showed changes in expression similar to those in the original analyses for eight out of nine snoRNAs (Supplementary Fig. 2), corroborating the results of the SOLiD analysis. In order to see if these results can be substantiated on a larger scale, we chose additional leukemic cell lines of T cell origin (Jurkat, P12-Ichikawa, Peer) and pre-B cell origin (Kasumi-2, Nalm-6, REH) (see Table 1). The expression of the nine DE snoRNAs was tested by RT-qPCR in the five T-ALL and five pre-B-ALL cell lines, and as shown in Fig. 3, six out of nine snoRNAs showed changes in expression similar to those in the original SOLiD data set. Among these were four box C/D snoRNAs (SNORD24, SNORD44, SNORD82, SNORD105) and two scaRNAs (scaRNA6 and scaRNA9). Statistically significant differences in expression between the pre-B-ALL and T-ALL subgroups could be seen in cases of scaRNA9 and SNORD24. It is worth noting that there was a marked variation in the expression of individual snoRNAs even within the leukemia subgroups, probably reflecting the unique characteristics of individual cell lines. We also analyzed the expression of the six snoRNAs by RT-qPCR using cDNA derived from

Table 2 DE snoRNAs selected for RT-qPCR analysis from T vs. pre-B comparison

name	alt name	fRNAdb accession	FC	p-value
scaRNA6	U88	FR022918, FR015998	3.284	0.0015
scaRNA9	mgU2-19/30, Z32	FR391181	1.938	0.0039
SNORD24	U24	FR017664	−1.485	0.1166
SNORD44	U44	FR355798	−1.826	0.0597
SNORD49A	U49A	FR191731	1.965	0.0053
SNORD55	U55	FR080434	1.723	9.43e-05
SNORD82	U82	FR223959	−1.531	0.1004
SNORD105	U105	FR045105	1.513	0.0051
SNORD110	HBII-55	FR358878	−1.737	0.0490

alt name alternative name; *fRNAdb accession* accession number in the Functional RNA database (http://www.ncRNA.org/frnadb); *FC* fold change

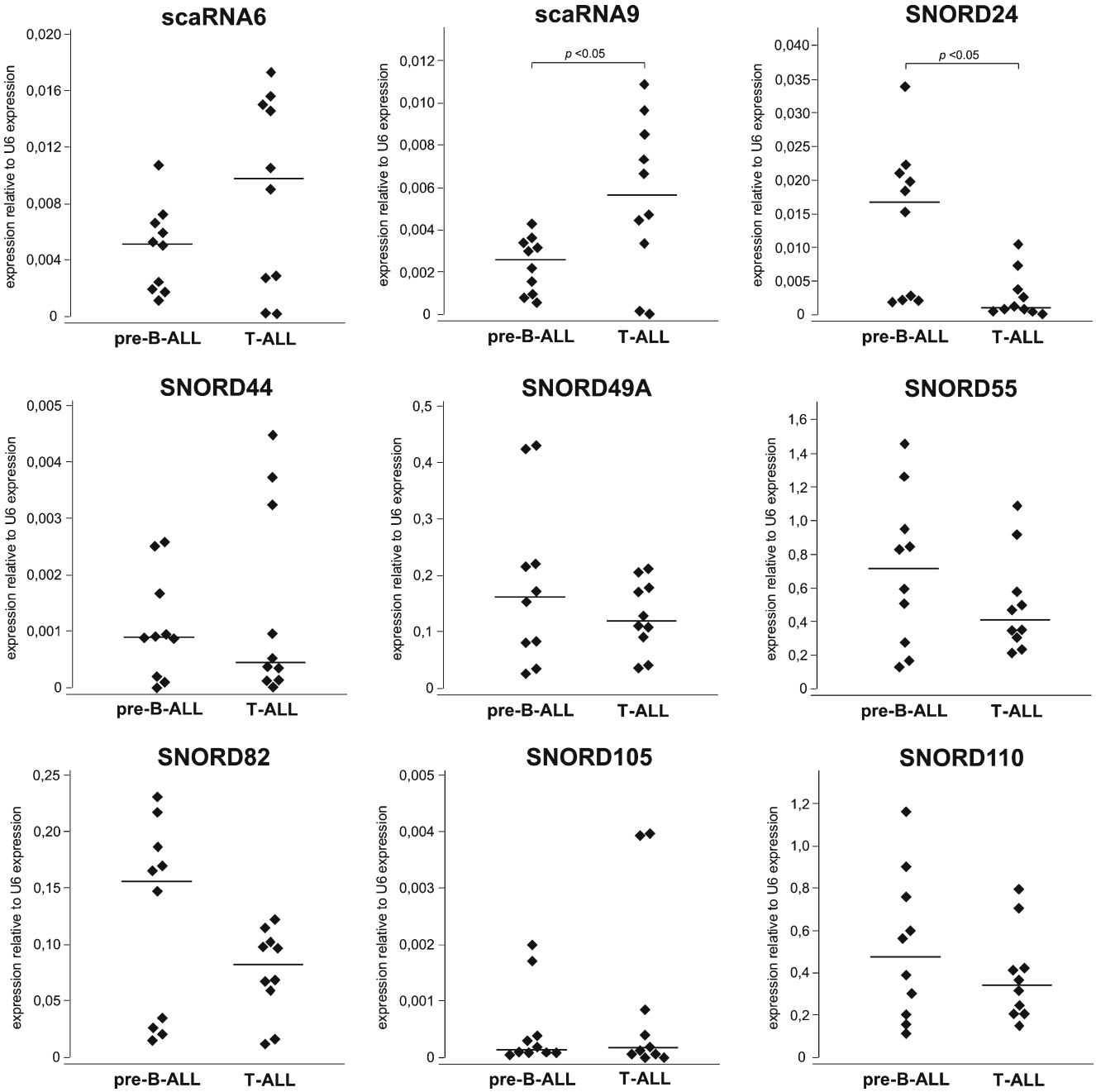

Fig. 3 Scatter plots exhibiting quantitative real-time PCR expression analyses of scaRNA6, scaRNA9, SNORD24, SNORD44, SNORD49A, SNORD55, SNORD82, SNORD105, and SNORD110 in five pre-B-ALL and five T-ALL cell lines. The pre-B-ALL cell lines were Kasumi-2, KOPN-8, MHH-CALL3, Nalm-6 and REH. The T-ALL cell lines were CCRF-CEM, Jurkat, Molt-16, P12-Ichikawa and Peer. Results from two independent runs are shown. The solid horizontal line represents the median value

total RNA (Supplementary Fig. 3). While scaRNA6 and scaRNA9 showed similar differences in expression, SNORD105 showed opposite difference in expression, and SNORD24, SNORD44 and SNORD82 were expressed at a very low level, making the interpretation unreliable. These latter results suggest that some low-expressed snoRNAs may be missed when using total RNA as a template.

Only a few reports discussing these snoRNAs have been published, and detailed descriptions are available for only SNORD24 [27] and SNORD82 [28]. SNORD44, encoded in the multi-snoRNA host gene Gas5 (growth arrest-specific transcript 5) [18], has been studied in cancer patients and a low expression of this snoRNA was shown to correlate with a poor prognosis [26]. Gas5 was also found to play an

important role in normal and transformed T-lymphocytes [29]. scaRNAs are known to be involved in post-transcriptional modification of snRNAs [4, 5].

While most snoRNAs are located in introns and co-transcribed with the host gene transcripts, a small number appears to be under the control of independent promoters [20]. Expression of snoRNAs may also be regulated by differential splicing of their host introns, since removal and disbranching of introns is essential for snoRNA production.

Valleron et al. [30] recently reported a global down-regulation of snoRNA expression in acute leukemias as compared to healthy blood samples. Specific snoRNA signatures were identified in leukemias of different subgroups, a finding that mirrors our results. Furthermore, they identified a locus containing SNORD112-114 as an ectopic expression cluster specific to acute promyeloid leukemias (APL). Expression of SNORD112-114 was found to be regulated by the APL-associated fusion gene PML-RARalpha, but the expression of individual snoRNAs in the cluster was independently regulated. Another recent report described down-regulation of a set of H/ACA box snoRNAs by the growth suppressor GRIM-1, which led to disruption of normal rRNA maturation [31].

There are only a handful of reports dealing with the association of snoRNAs and cancer. The U50 snoRNA has been implicated in prostate and breast cancers [32, 33]. In prostate cancer, a specific mutation in U50 was found to be associated with clinical manifestation of the disease [32]. In surgical specimens of non-small cell lung cancer (NSCLC), six snoRNAs were found to be overexpressed in the tumor tissue [34]. One of them, SNORA42, was reported to act as an oncogene in NSCLC, and its high expression in clinical samples correlated with poor survival [35]. Mourtada-Maarabouni et al. [36] discovered that Gas5, an RNA transcript hosting 10 snoRNAs (see above), regulated growth and apoptosis in mammalian cell lines, and that levels of the Gas5 transcript were significantly reduced in breast cancer samples.

Two classes of small ncRNA, miRNAs and snoRNAs, have both been associated with cancer. However, there are marked differences in the mechanisms of action and the molecular sizes between snoRNAs and miRNAs. While snoRNAs mainly direct modifications of rRNA, miRNAs usually bind to the 3′ untranslated region (3′-UTR) of a mRNA, resulting in its degradation or a block in its translation. Mature miRNA molecules are very short (~22 nucleotides), whereas C/D box snoRNAs are approximately 60–200 nucleotides, and H/ACA box snoRNAs 120–250 nucleotides, in length [22]. Nevertheless, similarities have also emerged between these two classes of ncRNA. Some snoRNAs have been shown to undergo further processing into smaller molecules with miRNA-like properties [23, 37], and some miRNA precursors have been found to resemble snoRNAs in their mode of action [38]. The involvement of miRNAs in normal as well as aberrant hematopoiesis is now well established [39, 40]. Several reports of miRNA profiling from patient serum, plasma, saliva, and other tissues highlight the diagnostic potential of these RNAs in the detection and monitoring of small tumors or those located deep in tissues [41–47]. Of particular relevance to the present work is the finding that snoRNAs can be detected in plasma [34].

4 Conclusions

We have shown here that various leukemia-derived cell lines carry a specific snoRNA expression signature, and we demonstrated specific differences in snoRNA expression between ALL subgroups. These results indicate that snoRNA expression profiles can potentially be used for the identification of novel subgroups of leukemia, and also hint at the possible involvement of snoRNAs in pathogenesis of leukemia. This work provides an impetus for further expression profiling of snoRNAs in normal and malignant processes.

Acknowledgements We thank Mr. Jorma Kulmala for technical assistance. This work was supported by the Academy of Finland Research Council for Health (funding decision number 115260), the Foundation for Paediatric Research in Finland, the Finnish Medical Foundation, the Competitive Research Funding of Tampere University Hospital (grants 9J062, 9K073 and 9M052), the Nona and Kullervo Väre Foundation and the Päivikki and Sakari Sohlberg Foundation. The funding sources had no involvement in the study.

Conflict of interest The authors declare no conflict of interest.

References

1. J.S. Mattick, I.V. Makunin, Non-coding RNA. Hum Mol Genet **15**, R17–29 (2006)
2. R.J. Taft, K.C. Pang, T.R. Mercer, M. Dinger, J.S. Mattick, Non-coding RNAs: regulators of disease. J Pathol **220**, 126–139 (2010)
3. T. Kiss, Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell. **109**, 145–148 (2002)
4. B.E. Jady, T. Kiss, A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J. **20**, 541–551 (2001)
5. X. Darzacq, B.E. Jady, C. Verheggen, A.M. Kiss, E. Bertrand, T. Kiss, Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. EMBO J **21**, 2746–2756 (2002)
6. T. Kiss, E. Fayet-Lebaron, B.E. Jady, Box H/ACA small ribonucleoproteins. Mol Cell. **37**, 597–606 (2010)
7. R.D. Leverette, M.T. Andrews, E.S. Maxwell, Mouse U14 snRNA is a processed intron of the cognate hsc70 heat shock pre-messenger RNA. Cell. **71**, 1215–1221 (1992)
8. P. Fragapane, S. Prislei, A. Michienzi, E. Caffarelli, I. Bozzoni, A novel small nucleolar RNA (U16) is encoded inside a ribosomal protein intron and originates by processing of the pre-mRNA. EMBO J. **12**, 2921–2928 (1993)
9. T. Kiss, W. Filipowicz, Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene RCC1. EMBO J **12**, 2913–2920 (1993)

10. K.T. Tycowski, M.D. Shu, J.A. Steitz, A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. Genes Dev. **7**, 1176–1190 (1993)
11. M.P. Hoeppner, S. White, D.C. Jeffares, A.M. Poole, Evolutionarily stable association of intronic snoRNAs and microRNAs with their host genes. Genome Biol Evol. **1**, 420–428 (2009)
12. K.T. Tycowski, M.D. Shu, J.A. Steitz, A mammalian gene with introns instead of exons generating stable RNA products. Nature **379**, 464–466 (1996)
13. O. Meyuhas, Synthesis of the translational apparatus is regulated at the translational level. Eur J Biochem **267**, 6321–6330 (2000)
14. J. Cavaille, J.P. Bachellerie, Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure. Biochimie **78**, 443–456 (1996)
15. N.J. Watkins, R.D. Leverette, L. Xia, M.T. Andrews, E.S. Maxwell, Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. RNA. **2**, 118–133 (1996)
16. S.L. Ooi, D.A. Samarsky, M.J. Fournier, J.D. Boeke, Intronic snoRNA biosynthesis in saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. RNA. **4**, 1096–1110 (1998)
17. P. Pelczar, W. Filipowicz, The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5′-terminal oligopyrimidine gene family. Mol Cell Biol. **18**, 4509–4518 (1998)
18. C.M. Smith, J.A. Steitz, Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5′-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. Mol Cell Biol. **18**, 6897–6909 (1998)
19. M.E. Askarian-Amiri, J. Crawford, J.D. French, C.E. Smart, M.A. Smith, M.B. Clark, K. Ru, T.R. Mercer, E.R. Thompson, S.R. Lakhani, A.C. Vargas, I.G. Campbell, M.A. Brown, M.E. Dinger, J.S. Mattick, SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. RNA. **17**, 878–891 (2011)
20. G. Dieci, M. Preti, B. Montanini, Eukaryotic snoRNAs: a paradigm for gene expression flexibility. Genomics. **94**, 83–88 (2009)
21. R.J. Taft, E.A. Glazov, T. Lassmann, Y. Hayashizaki, P. Carninci, J.S. Mattick, Small RNAs derived from snoRNAs. RNA. **15**, 1233–1240 (2009)
22. M.S. Scott, M. Ono, From snoRNA to miRNA: dual function regulatory non-coding RNAs. Biochimie. **93**, 1987–1992 (2011)
23. M. Brameier, A. Herwig, R. Reinhardt, L. Walter, J. Gruber, Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic Acids Res. **39**, 675–686 (2011)
24. S. Anders, W. Huber, Differential expression analysis for sequence count data. Genome Biol. **11**, R106 (2010)
25. T.D. Schmittgen, J. Jiang, Q. Liu, L. Yang, A high-throughput method to monitor the expression of microRNA precursors. Nucleic Acids Res. **32**, e43 (2004)
26. H.E. Gee, F.M. Buffa, C. Camps, A. Ramachandran, R. Leek, M. Taylor, M. Patil, H. Sheldon, G. Betts, J. Homer, C. West, J. Ragoussis, A.L. Harris, The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. Br J Cancer. **104**, 1168–1177 (2011)
27. L.H. Qu, Y. Henry, M. Nicoloso, B. Michot, M.C. Azum, M.H. Renalier, M. Caizergues-Ferrer, J.P. Bachellerie, U24, a novel intron-encoded small nucleolar RNA with two 12 nt long, phylogenetically conserved complementarities to 28S rRNA. Nucleic Acids Res. **23**, 2669–2676 (1995)
28. A. Rebane, A. Metspalu, U82, a novel snoRNA identified from the fifth intron of human and mouse nucleolin gene. Biochim Biophys Acta. **1446**, 426–430 (1999)
29. G.T. Williams, M. Mourtada-Maarabouni, F. Farzaneh, A critical role for non-coding RNA GAS5 in growth arrest and rapamycin inhibition in human T-lymphocytes. Biochem Soc Trans. **39**, 482–486 (2011)
30. W. Valleron, E. Laprevotte, E.F. Gautier, C. Quelen, C. Demur, E. Delabesse, X. Agirre, F. Prósper, T. Kiss, P. Brousset, Specific small nucleolar RNA expression profiles in acute leukemia. Leukemia (2012). doi:10.1038/leu.2012.111
31. S.C. Nallar, L. Lin, V. Srivastava, P. Gade, E.R. Hofmann, H. Ahmed, S.P. Reddy, D.V. Kalvakolanu, GRIM-1, a novel growth suppressor, inhibits rRNA maturation by suppressing small nucleolar RNAs. PLoS One **6**, e24082 (2011)
32. X.Y. Dong, C. Rodriguez, P. Guo, X. Sun, J.T. Talbot, W. Zhou, J. Petros, Q. Li, R.L. Vessella, A.S. Kibel, V.L. Stevens, E.E. Calle, J.T. Dong, SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. Hum Mol Genet **17**, 1031–1042 (2008)
33. X.Y. Dong, P. Guo, J. Boyd, X. Sun, Q. Li, W. Zhou, J.T. Dong, Implication of snoRNA U50 in human breast cancer. J Genet Genomics. **36**, 447–454 (2009)
34. J. Liao, L. Yu, Y. Mei, M. Guarnera, J. Shen, R. Li, Z. Liu, F. Jiang, Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. Mol Cancer. **9**, 198 (2010)
35. Y.P. Mei, J.P. Liao, J.P. Shen, L. Yu, B.L. Liu, L. Liu, R.Y. Li, L. Ji, S.G. Dorsey, Z.R. Jiang, R.L. Katz, J.Y. Wang, F. Jiang, Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. Oncogene. **31**, 2794–2804 (2012)
36. M. Mourtada-Maarabouni, M.R. Pickard, V.L. Hedge, F. Farzaneh, G.T. Williams, GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene. **28**, 195–208 (2009)
37. C. Ender, A. Krek, M.R. Friedlander, M. Beitzinger, L. Weinmann, W. Chen, S. Pfeffer, N. Rajewsky, G. Meister, A human snoRNA with microRNA-like functions. Mol Cell. **32**, 519–528 (2008)
38. M. Ono, M.S. Scott, K. Yamada, F. Avolio, G.J. Barton, A.I. Lamond, Identification of human miRNA precursors that resemble box C/D snoRNAs. Nucleic Acids Res. **39**, 3879–3891 (2011)
39. V. Havelange, R. Garzon, MicroRNAs: emerging key regulators of hematopoiesis. Am J Hematol. **85**, 935–942 (2010)
40. S. Bhagavathi, M. Czader, MicroRNAs in benign and malignant hematopoiesis. Arch Pathol Lab Med. **134**, 1276–1281 (2010)
41. P.S. Mitchell, R.K. Parkin, E.M. Kroh, B.R. Fritz, S.K. Wyman, E.L. Pogosova-Agadjanyan, A. Peterson, J. Noteboom, K.C. O'Briant, A. Allen, D.W. Lin, N. Urban, C.W. Drescher, B.S. Knudsen, D.L. Stirewalt, R. Gentleman, R.L. Vessella, P.S. Nelson, D.B. Martin, M. Tewari, Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A **105**, 10513–10518 (2008)
42. X. Chen, Y. Ba, L. Ma, X. Cai, Y. Yin, K. Wang, J. Guo, Y. Zhang, J. Chen, X. Guo, Q. Li, X. Li, W. Wang, Y. Zhang, J. Wang, X. Jiang, Y. Xiang, C. Xu, P. Zheng, J. Zhang, R. Li, H. Zhang, X. Shang, T. Gong, G. Ning, J. Wang, K. Zen, J. Zhang, C.Y. Zhang, Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res. **18**, 997–1006 (2008)
43. E.K. Ng, W.W. Chong, H. Jin, E.K. Lam, V.Y. Shin, J. Yu, T.C. Poon, S.S. Ng, J.J. Sung, Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. Gut **58**, 1375–1381 (2009)
44. M.A. Cortez, G.A. Calin, MicroRNA identification in plasma and serum: a new tool to diagnose and monitor diseases. Expert Opin Biol Ther. **9**, 703–711 (2009)

45. M.I. Aslam, K. Taylor, J.H. Pringle, J.S. Jameson, MicroRNAs are novel biomarkers of colorectal cancer. Br J Surg. **96**, 702–710 (2009)
46. M. Tsujiura, D. Ichikawa, S. Komatsu, A. Shiozaki, H. Takeshita, T. Kosuga, H. Konishi, R. Morimura, K. Deguchi, H. Fujiwara, K. Okamoto, E. Otsuji, Circulating microRNAs in plasma of patients with gastric cancers. Br J Cancer. **102**, 1174–1179 (2010)
47. Z. Hu, X. Chen, Y. Zhao, T. Tian, G. Jin, Y. Shu, Y. Chen, L. Xu, K. Zen, C. Zhang, H. Shen, Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. J Clin Oncol. **28**, 1721–1726 (2010)

Authors' contributions

Kaisa J. Teittinen and Olli Lohi provided conception of the study, designed the study and drafted the manuscript. Kaisa J. Teittinen, Asta Laiho, Annemari Uusimäki and Juha-Pekka Pursiheimo performed the experiments. Kaisa J. Teittinen, Asta Laiho, Attila Gyenesei and Olli Lohi analyzed and interpreted the data. Asta Laiho, Annemari Uusimäki, Juha-Pekka Pursiheimo and Attila Gyenesei revised the manuscript and all authors approved the final version of the submitted manuscript.

http://www.nature.com/leu/journal/vaop/ncurrent/full/leu2013250a.html

Leukemia (2013), 1–3

www.nature.com/leu

LETTER TO THE EDITOR

Overexpression of SNORD114-3 marks acute promyelocytic leukemia

Leukemia advance online publication, 20 September 2013; doi:10.1038/leu.2013.250

A global downregulation of small nucleolar RNAs (snoRNA) was reported in acute lymphoblastic (ALL) and acute myeloid leukemias (AML) by Valleron *et al.*[1] The authors also identified a specific snoRNA signature in acute promyelocytic leukemia (APL) with overexpression of snoRNA clusters *SNORD112-114*. Recently, Cohen *et al.*[2] reported a marked overexpression of *SNORD113-3*, *SNORD113-4*, *SNORD114-2* and *SNORD114-3* in three additional cases of APL.

We have compiled a hematological gene expression data set of neoplastic and non-neoplastic samples hybridized to Affymetrix HG-U133 Plus 2.0 GeneChips, all downloaded from the GEO microarray repository. The data were collected as described earlier by Heinäniemi *et al.*[3] Of all the genes on the microarray, 15 were identified as snoRNAs (Table 1). They represent various types of snoRNAs, including H/ACA and CD box types, small Cajal body-specific RNAs (scaRNA), intergenic and intronic snoRNAs, and even one snoRNA without any known target RNA (so-called 'orphan' snoRNA).

To study the expression of the 15 snoRNAs, we focused on 883 cases of pediatric leukemias, which were subdivided into three classes: early B-ALL, T-ALL and AML. In addition, we considered 35 non-neoplastic samples consisting of hematopoietic stem cells (HSCs) along with naive B- and naive T-lymphocytes. In most cases, the snoRNAs were expressed in a relatively uniform manner across all leukemia subtypes and normal blood cells (Figure 1a). In our limited set of 15 snoRNAs, we did not observe a lower expression in acute leukemias as compared with HSCs and naive lymphocytes. Two snoRNAs, *SNORA70* and *SNORD104*, were expressed consistently higher than other snoRNAs in leukemic as well as healthy samples. They were expressed at a level approximately fourfold higher relative to other snoRNAs (Figure 1a). Interestingly, *SNORA25* and *SNORA61* were expressed strongly in naive, or unstimulated, B cells compared to HSCs, naive T cells and leukemias, suggesting B-cell differentiation-dependent regulation of expression (Figure 1a).

At single patient level, the expression of individual snoRNAs remained again rather constant except for *SNORD114-3* (Figure 1a and Supplementary Figure S1). Among the 237 pediatric AML patients, 15 were found to have increased expression of *SNORD114-3* when the cut-off level in $\log_2$-expression was set at 7 (Figure 1b). Interestingly, 13 out of these 15 patients harbored t(15;17) translocation, the hallmark of APL. In the two remaining t(15;17)-negative patients, the expression was only slightly above the threshold level of 7. These two samples did not fall into any major cytogenetic subtype of AML, but one of them had an internal tandem duplication of FLT3 and the other carried a mutation in either *NRAS* or *KRAS*. Among the cases with normal *SNORD114-3* expression, only six patients out of 222, or 2.7%, were identified as t(15;17)-positive patients. Four of these six APL cases were carrying FLT3-ITD. A similar pattern emerged also in the cohort of 1117 adult AML patients in our database, 69 of which were of promyelocytic subtype (Figure 1c). One-hundred and nine adult AML samples were found with an elevated *SNORD114*-3 expression (keeping the threshold at 7). Out of them, 50 were from APL patients. From the adult samples with normal *SNORD114-3* expression, only 19, or 1.9%, were classified as APL. The expression level of *SNORD114-3* differentiated the APL-positive and -negative cases among both pediatric ($P < 10^{-8}$) and adult AML ($P < 10^{-28}$). Mann–Whitney *U*-test was used to assess statistical significance.

Table 1. Characteristics of 15 selected snoRNAs

snoRNA	*Type*	*Host gene*	*Chromosome*	*Target*	*Probes*
SCARNA2	Small cajal body	Independent transcriptional unit	1q13.1	U2 snRNA C61 and U2 snRNA G11	11
SCARNA17	Small cajal body	Independent transcriptional unit	18q21.1	U4 snRNA C8	11
SNORA25	Box H/ACA	*Homo sapiens* cDNA AK128061	11q21	18S rRNA U801 and 18S rRNA U814	6
SNORA61	Box H/ACA	SNHG12 small nucleolar RNA host gene 12 (non-protein coding)	1p35.3	28S rRNA U2495	7
SNORA64	Box H/ACA	RPS2 (ribosomal protein S2)	16p13.3	28S rRNA U4975	10
SNORA65	Box H/ACA	RPL12 (ribosomal protein L12)	9q34	28S rRNA U4373 and 28S rRNA U4427	4
SNORA66	Box H/ACA	RPL5 (ribosomal protein L5)	1p22.1	18S rRNA U119	5
SNORA68	Box H/ACA	RPL18A (ribosomal protein L18a)	19p13	28S rRNA U4393	22
SNORA70	Box H/ACA	RPL10 (ribosomal protein L10)	Xq28	18S rRNA U1692	4
SNORA71A	Box H/ACA	SNHG17 small nucleolar RNA host gene 17 (non-protein coding)	20q11.23	18S rRNA U406	4
SNORA71B	Box H/ACA	SNHG17 small nucleolar RNA host gene 17 (non-protein coding)	20q11.23	18S rRNA U406	22
SNORA74A	Box H/ACA	SNHG4 small nucleolar RNA host gene 4; MATR3 alt Matrin3 mRNA	5q31.2	28S rRNA U3741, 28S rRNA U3743 and U3 snRNA U8	33
SNORD8	Box C/D	CHD8 chromodomain helicase DNA binding protein 8	14q11.2	U6 snRNA A53	5
SNORD104	Box C/D	Independent transcriptional unit	17q23.3	28S rRNA C1327	5
SNORD114-3	Box C/D	LOC100507242, snoRNA gene cluster	14q32	Unknown	4

Accepted article preview online 27 August 2013

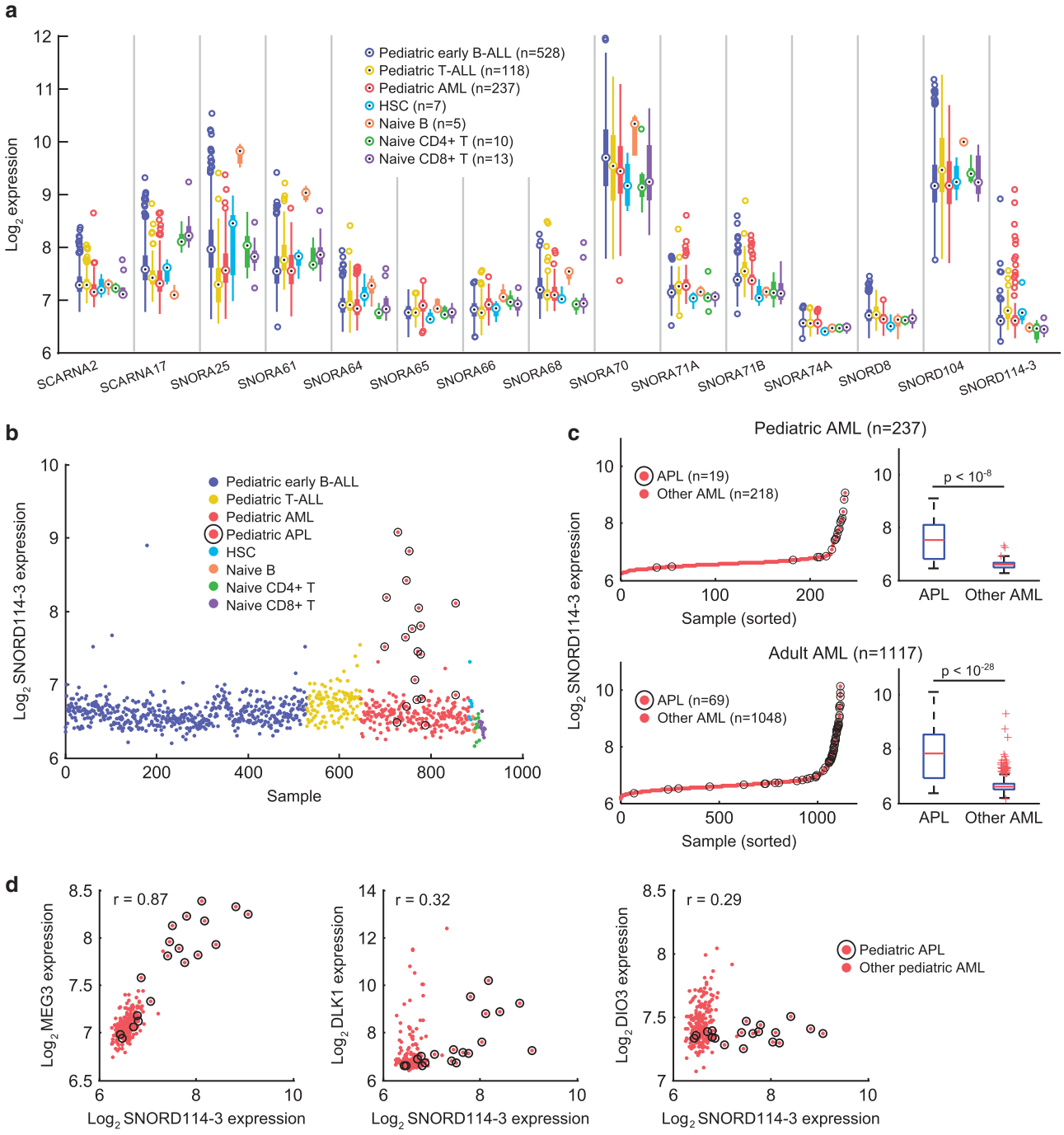

Figure 1. Overexpression of *SNORD114-3* in APL. (**a**) Expression of the 15 studied snoRNAs in three types of pediatric leukemias and four healthy cell types. (**b**) The expression of *SNORD114-3* in pediatric leukemias. (**c**) When pediatric and adult AML patients are sorted according to *SNORD114-3* expression, APL cases are clearly enriched among the patients with high expression. By using Mann–Whitney *U*-test, differential *SNORD114-3* expression between APL and other AML was deemed statistically significant. (**d**) *SNORD114-3* expression is strongly correlated with *MEG3* expression, but weakly with *DLK1* and *DIO3*. Pearson's correlation coefficients were calculated using $\log_2$ expressions.

The snoRNA cluster *SNORD114* lies in the genomically imprinted domain DLK1-DIO3 in the 14q32 region. Genomic imprinting gives rise to mono-allelic, parent-of-origin-specific gene expression. Imprinted genes are susceptible to errors, as a single genomic or epigenetic change may alter or even ablate their function. Imprinted small RNAs are involved in the development and metabolism, and are also frequently perturbed in malignancies.[4] The paternal allele of DLK1-DIO3 domain expresses three protein-coding genes (*DLK1*, *RTL1* and *DIO3*), whereas the maternal counterpart expresses two non-coding transcripts (*MEG3* and *RTL1AS*) accompanied by 53 microRNAs and two snoRNA clusters (*SNORD113* and *-114*) constituting 9 and 31 snoRNAs, respectively.[4] To gain some insight on how deregulation of DLK1-DIO3 affects other genes it houses, we looked at the expressions of *DLK1*, *DIO3* and *MEG3*. Besides *SNORD114-3*, these were the only genes in this domain to be represented on the microarray. The expression of *MEG3* (*Maternally Expressed Gene 3*) was strongly correlated with that of *SNORD114-3*, whereas such correlation was

not observed between *SNORD114-3* and *DLK1* or *DIO3* (Figure 1d). This reflects the fact that *MEG3* is expressed from the same (maternal) allele as *SNORD114*, whereas *DLK1* and *DIO3* are expressed from the other (paternal) one. Furthermore, it suggests that the deregulation may affect the maternal DLK1-DIO3 locus as a whole.

As it happens, *SNORD114-3*, like all the copies in *SNORD114*, is orphan. The function of such snoRNAs lacking RNA targets is unknown but they might have gene regulatory roles as microRNA precursors or by being involved in gene splicing events.[5] Valleron *et al.*[1] showed that *SNORD114-1* is capable of modulating cell growth and proliferation in APL cells, thus suggesting leukemogenicity of *SNORD114*. Recently, Chu *et al.*[6] found that *ACA11*, an orphan snoRNA encoded in an intron of the *WHSC1* gene, is overexpressed in multiple myeloma patients with the translocation t(4;14). In these cells, overexpression of *ACA11* increased proliferation, suppressed oxidative stress and conferred chemoresistance.

Our findings in regard to APL support and increase the robustness of those reported by Valleron *et al.*[1] and Cohen *et al.*,[2] as we used data obtained with a different platform and from a substantially larger cohort of patients. Collectively, they reveal an intricate regulation of snoRNA expression[7] and add to the mounting evidence implicating DLK1-DIO3 locus in tumorigenesis. Seven microRNAs of this locus were found to be upregulated in APL by Dixon-McIver *et al.*,[8] and other transcripts of DLK1-DIO3 domain have been associated with lung cancer,[9] hepatocellular carcinoma[10] and the pluripotency levels of stem cells.[11]

CONFLICT OF INTEREST

The authors declare no conflict of interest.

T Liuksiala[1,2], KJ Teittinen[3], K Granberg[1,2], M Heinäniemi[4], M Annala[1,2], M Mäki[3], M Nykter[1,5] and O Lohi[3]
[1]*Institute of Biomedical Technology, University of Tampere, Tampere, Finland;*
[2]*Department of Signal Processing, Tampere University of Technology, Tampere, Finland;*
[3]*Tampere Center for Child Health Research, University of Tampere and Tampere University Hospital, Tampere, Finland;*
[4]*A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland and*
[5]*BioMediTech, Tampere, Finland*
E-mail: olli.lohi@uta.fi

REFERENCES

1 Valleron W, Laprevotte E, Gautier EF, Quelen C, Demur C, Delabesse E *et al.* Specific small nucleolar RNA expression profiles in acute leukemia. *Leukemia* 2012; **26**: 2052–2060.
2 Cohen Y, Hertzog K, Reish O, Mashevich M, Garach-Jehoshua O, Bar-Chaim A *et al.* The increased expression of 14q32 small nucleolar RNA transcripts in promyelocytic leukemia cells is not dependent on PML-RARA fusion gene. *Blood Cancer J* 2012; **2**: e92.
3 Heinäniemi M, Nykter M, Kramer R, Wienecke-Baldacchino A, Sinkkonen L, Zhou JX *et al.* Gene-pair expression signatures reveal lineage control. *Nat Methods* 2013; **10**: 577–583.
4 Girardot M, Cavaillé J, Feil R. Small regulatory RNAs controlled by genomic imprinting and their contribution to human disease. *Epigenetics* 2012; **7**: 1341–1348.
5 Taulli R, Pandolfi PP. 'Snorkeling' for missing players in cancer. *J Clin Invest* 2012; **122**: 2765–2768.
6 Chu L, Su MY, Maggi Jr LB, Lu L, Mullins C, Crosby S *et al.* Multiple myeloma-associated chromosomal translocation activates orphan snoRNA ACA11 to suppress oxidative stress. *J Clin Invest* 2012; **122**: 2793–2806.
7 Teittinen KJ, Laiho A, Uusimäki A, Pursiheimo JP, Gyenesei A, Lohi O. Expression of small nucleolar RNAs in leukemic cells. *Cell Oncol* 2013; **36**: 55–63.
8 Dixon-McIver A, East P, Mein CA, Cazier JB, Molloy G, Chaplin T *et al.* Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia. *PLoS One* 2008; **3**: e2141.
9 Liu X, Sempere LF, Ouyang H, Memoli VA, Andrew AS, Luo Y *et al.* MicroRNA-31 functions as an oncogenic microRNA in mouse and human lung cancer cells by repressing specific tumor suppressors. *J Clin Invest* 2010; **4**: 1298–1309.
10 Wang PR, Xu M, Toffanin S, Li Y, Llovet JM, Russell DW. Induction of hepatocellular carcinoma by *in vivo* gene targeting. *Proc Natl Acad Sci USA* 2012; **109**: 11264–11269.
11 Liu L, Luo GZ, Yang W, Zhao X, Zheng Q, Lv Z *et al.* Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J Biol Chem* 2010; **285**: 19483–19490.

Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)

Leukemia Research 36 (2012) 232–236

Contents lists available at ScienceDirect

Leukemia Research

journal homepage: www.elsevier.com/locate/leukres

Nucleolar proteins with altered expression in leukemic cell lines

Kaisa J. Teittinen[a,*], Pauliina Kärkkäinen[a], Johanna Salonen[b], Gunilla Rönnholm[c], Hanna Korkeamäki[a], Mauno Vihinen[b], Nisse Kalkkinen[c], Olli Lohi[a]

[a] Paediatric Research Centre, University of Tampere Medical School and Tampere University Hospital, Tampere, Finland
[b] Institute of Biomedical Technology, University of Tampere, Tampere, Finland
[c] Institute of Biotechnology, University of Helsinki, Helsinki, Finland

ARTICLE INFO

Article history:
Received 12 May 2011
Received in revised form 28 June 2011
Accepted 29 June 2011
Available online 23 July 2011

Keywords:
Leukemia
Nucleolus
Proteomics
Mass spectrometry
Cell line
Differential protein expression

ABSTRACT

The function of the nucleolus is intimately connected to cell proliferation, division and growth. Many cancer cells have enlarged nucleoli, and several nucleolar proteins have been linked to tumorigenesis. In order to find proteins whose expression is altered in the nucleoli of leukemic cells, we carried out two-dimensional difference gel electrophoresis (2-D DIGE) analyses. Prohibitin (PHB) and TAR-DNA-binding protein-43 (TDP-43) were strongly expressed in the nucleoli of the pre-B-ALL cell line MHH-CALL3. Our results demonstrate that leukemic cells have differences in their nucleolar protein composition, and suggest that it may be possible to exploit these differences in identification of leukemia subtypes.

1. Introduction

The nucleolus is a prominent nuclear structure best known for its central role in ribosome biogenesis. When a cell enters mitosis, the nucleoli are disassembled, and at the end of mitosis they begin to reassemble around clusters of ribosomal RNA genes, called nucleolar organizer regions (NORs). The nucleolus, which is not bound by a membrane, can be divided into 3 sub-regions: the fibrillar centre (FC), the dense fibrillar component (DFC) and the granular component (GC) [1–3]. In addition to roles in ribosome biogenesis, other functions have been identified for many nucleolar proteins. Functions ascribed to the nucleolus include control of the cell-cycle, cellular stress responses and synthesis of small RNAs and RNA–protein complexes [4]. Over 4500 proteins are believed to localize to the nucleolus in human cells and these proteins are estimated to represent over 80% of the total nucleolar proteome [5]. The nucleolus appears to be highly conserved through eukaryotic evolution, since approximately 90% of human nucleolar proteins have homologues in yeast [6]. The size and number of nucleoli in human cells vary according to the cell type and the proliferation rate of the cell.

Consistent with the nucleolus having a significant influence on cell proliferation and growth, disturbances and alterations in its structure and function have been identified in a number of diseases, including cancer. For instance, inherited diseases such as Diamond-Blackfan anemia, dyskeratosis congenita and Shwachman-Diamond syndrome, which are also associated with increased risk of cancer, are caused by mutated nucleolar proteins [7]. A nucleolar protein called nucleophosmin has been found to be mutated or overexpressed in various tumors and hematologic malignancies [8]. In acute myeloid leukemia with no karyotypic changes, mutated nucleophosmin mislocalizes to the cytoplasm instead of the nucleolus in more than 30% of the patients [9].

Encouraged by these discoveries, this study was undertaken to find changes in nucleolar proteomes in various acute leukemic cell lines. The aim was to identify nucleolar proteins that could be used as markers in the classification or diagnostics of various subtypes of leukemia. Here we provide evidence that prohibitin (PHB) and TAR DNA-binding protein 43 (TDP-43) are differentially expressed in the nucleoli of leukemic cell lines.

2. Materials and methods

2.1. Cell culture

The cell lines used were: MHH-CALL3, KOPN-8, CCRF-CEM and MOLT-16 (acute lymphatic leukemia, ALL), OCI-AML3 and HL-60 (acute myeloid leukemia, AML), and HeLa (cervical carcinoma). KOPN-8, CCRF-CEM, MOLT-16 and HL-60 cells were cultured in RPMI-1640 medium (Gibco) supplemented with 1% penicillin/streptomycin

* Corresponding author at: Paediatric Research Centre, University of Tampere Medical School, FM-3, Biokatu 10, 3rd Floor, FIN-33520 Tampere, Finland. Tel.: +358 3 3551 8406; fax: +358 3 3551 8402.
E-mail address: kaisa.teittinen@uta.fi (K.J. Teittinen).

0145-2126/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.leukres.2011.06.038

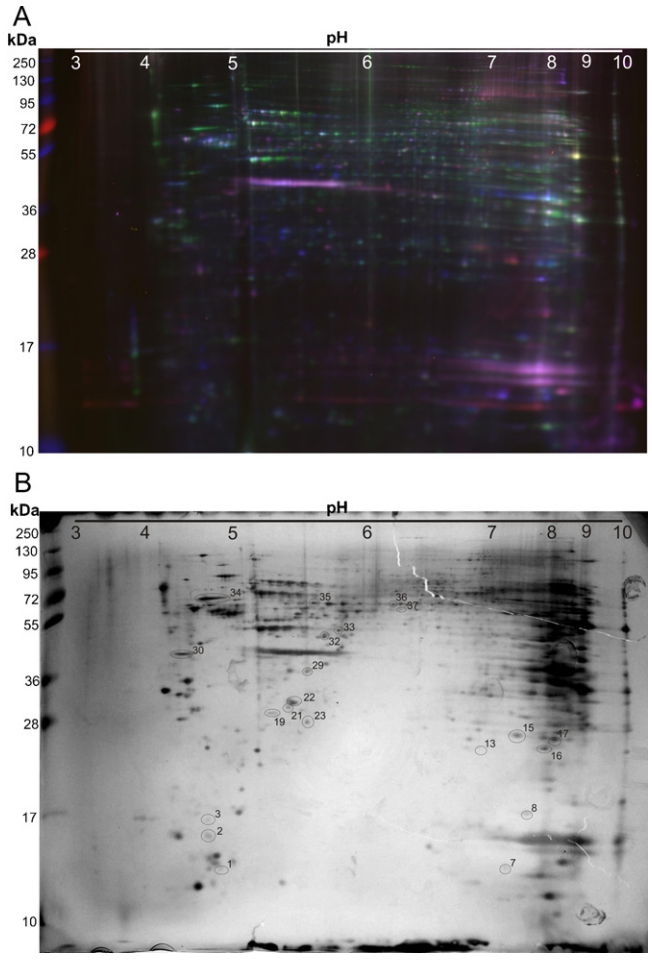

Fig. 1. 2-D DIGE analysis of nucleolar extracts shows differentially expressed protein spots between a pre-B-ALL and a pre-T-ALL cell line. (A) A representative 2D-DIGE fluorescence gel. (B) The same gel stained with silver nitrate. Molecular weights and pH gradient as indicated. The protein spots excised from the gel are marked (see Table 1).

antibiotics (pen-strep), 1% L-glutamine (L-glu) and 10% fetal bovine serum (FBS). MHH-CALL3 cells were cultured with the same medium, but with 20% FBS. OCI-AML3 cells were cultured in MEM Alpha Medium (Gibco) supplemented with pen-strep, L-glu and 20% FBS. HeLa cells were cultured in D-MEM medium (Sigma) supplemented with pen-strep, L-glu, non-essential amino acids and 10% FBS.

2.2. Isolation of nucleoli

For the isolation of nucleoli, a cell fractionation protocol by Andersen et al. [10] was used with some modifications. Briefly, the leukemic cells were grown to about 90% confluency and 1–3 × 10^8 cells were collected by centrifugation. The cells were washed with PBS and resuspended in 5 ml of a hypotonic buffer (10 mM Hepes (pH 7.9), 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT), broken on ice by ten strokes in a dounce homogenizer with a tight pestle, after which the nuclei were collected by centrifugation (228 × *g*, 5 min, +4 °C). To remove impurities, the nuclear pellet was resuspended in 3 ml of S1-solution (0.35 M sucrose, 10 mM $MgCl_2$), layered over 3 ml of S2-solution (0.35 M sucrose, 0.5 mM $MgCl_2$), and centrifuged (1430 × *g*, 5 min, +4 °C). The pelleted nuclei were then resuspended in 3 ml of S2-solution and sonicated for 3–5 min on ice using a 10 s/20 s ON/OFF-cycle at power setting "high" (Bioruptor UCD-200, Diagenode). The sonicate was checked for the absence of intact nuclei and presence of nucleoli with a phase contrast microscope (Axiovert 200 M, Zeiss). To pellet the nucleoli, the sonicate was layered over 3 ml of S3-solution (0.88 M sucrose, 0.5 mM $MgCl_2$) and centrifuged (2800 × *g*, 10 min, +4 °C). The pelleted nucleoli were further washed in 500 μl of S2-solution and centrifuged (2800 × *g*, 5 min, +4 °C). The purified nucleoli were resuspended in 500 μl of S2-solution and stored at −70 °C until use. The nucleolar content and quality of the samples were assessed by using immunofluorescence microscopy with the anti-nucleophosmin (32-5200, ZYMED) antibody and DAPI staining (for DNA). The total protein concentration of the nucleolar extracts was determined using a standard protein assay (DC Protein Assay, Bio-Rad).

2.3. Two-dimensional difference gel electrophoresis (2-D DIGE)

2-D DIGE is comprised of two dimensions: isoelectric focusing (IEF), which separates proteins according to isoelectric point (pI), and SDS-PAGE, which separates proteins according to molecular weight. For 2-D DIGE, the Ettan DIGE system (GE Healthcare) was used according to the manufacturer's instructions. The nucleolar samples were from CCRF-CEM, MHH-CALL3, OCI-AML3 or HeLa cell lines. For each sample, 3 independent nucleolar preparations were pooled. The pooled internal standard was comprised of nucleolar and nuclear extracts of CCRF-CEM, MHH-CALL3, OCI-AML3 and HeLa cell lines in equal amounts. The samples were labeled with fluorescent cyanide minimal dyes (CyDye DIGE Fluor dyes, GE Healthcare) according to the manufacturer's instructions. Each gel contained two samples (labeled with CyDye3 or CyDye5) and an internal standard (labeled with CyDye2), totaling 150 μg of protein.

IEF was performed using an Ettan IPGphor II Isoelectric Focusing System (GE Healthcare). The samples in rehydration buffer were loaded onto 24-cm non-linear immobilized pH gradient (IPG) strips (pH 3–10), and rehydration was performed at 20 °C for 6 h followed by 5.5 h at a low voltage (20 V). The current limit was set to 50 μA per IPG strip. IEF was carried out for a total of 45–47 kV h at 20 °C in 6 steps which gradually increased the voltage (180 V for 0.75 h, 300 V for 0.75 h, 600 V for 0.75 h, gradient from 600 V to 5000 V for 1 h, gradient from 5000 V to 8000 V for 0.75 h, and 8000 V about 5 h). After IEF, the IPG strips were stored at −20 °C.

Prior to SDS-PAGE the IPG strips were equilibrated for 15 min in SDS-equilibrium buffer (0.75 M Tris–HCl, pH 8.8, 6 M urea, 30% glycerol, 2% SDS, 10 mg/ml DTT and a trace of bromophenol blue), and then for 15 min in the same buffer containing 25 mg/ml iodoacetamide instead of DTT. The IPG strips were placed on top of 12% SDS-PAGE gels, a molecular weight marker was added next to the strip, and the gels were run at 1-2 W per gel for approximately 12 h at 20 °C using the Ettan DALTsix system (GE Healthcare). In order to facilitate spot picking, the gels were post-stained with silver using a method compatible with identification of proteins by mass-spectrometry (MS) [11].

2.4. Gel imaging and protein identification

CyDye-labeled protein spots were visualized with a Typhoon™ Trio Variable Mode Imager (GE Healthcare) at the appropriate wavelengths. Analysis of the spot map images was performed using DeCyder™ v6.5 (GE Healthcare) according to the manufacturer's recommendations, and *p*-values were calculated by Student's test. Protein spots of interest were cut from silver-stained 2-D gels and digested 'in-gel' essentially as described by Rosenfeld et al. [12] and Wilm et al. [13]. Proteins were reduced with DTT and alkylated with iodoacetamide before digestion with trypsin (Sequencing Grade Modified Trypsin, Promega Corporation, Madison, WI, USA) at 37 °C overnight. The peptides were extracted once with 25 mM ammonium bicarbonate and twice with 5% formic acid for 15 min at room temperature, after which the extracts were pooled. The peptides were analyzed using an automated nanoLC–MS/MS that consisted of a QSTAR® Elite Hybrid LC/MS/MS System (Applied Biosystems) coupled to a nanoLC instrument UltiMate® 3000 (Dionex Corporation). Database searches with the obtained MS/MS data were performed with the MASCOT MS/MS Ion Search program (http://www.matrixscience.com) against the MSDB database. The probability of a correct match to a known protein was estimated by considering the protein's MASCOT score, the number of peptide matches and sequence coverage, as well as the correspondence of the experimental and theoretical molecular weight and pI values. According to the peptide summary report of the MASCOT search program, ion scores above 36 for individual peptides could be considered significant ($p < 0.05$), and the MASCOT score represents the sum of ion scores for each protein. Sequence coverage was calculated from the number of amino acids with matches in a peptide, as a percentage of the entire amino acid sequence of the protein. The more peptides that could be matched to a protein, especially with significant ion scores, and the higher the sequence coverage and the MASCOT score, the more probable was a correct match.

2.5. Western blotting and immunofluorescence

Samples of the nucleolar preparations were prepared in 2× Laemmli buffer, and the protein concentrations were adjusted to 0.5 mg/ml. Proteins were resolved by SDS-PAGE and transferred to nitrocellulose membranes (Amersham Biosciences). The immunoblots were probed with anti-PHB (Santa Cruz Biotechnology) or anti-TDP-43 (TDP-43 and TDP-43-N, Sigma–Aldrich) antibodies, and the proteins were detected by enhanced chemiluminescence. The same antibodies were used for immunofluorescence staining of the nucleolar preparations. DAPI was used for visualization of DNA. AlexaFluor-conjugated secondary antibodies were used (Molecular Probes, Invitrogen).

3. Results

In order to identify nucleolar proteins involved in leukemogenesis, we used a proteomic approach to look for proteins whose

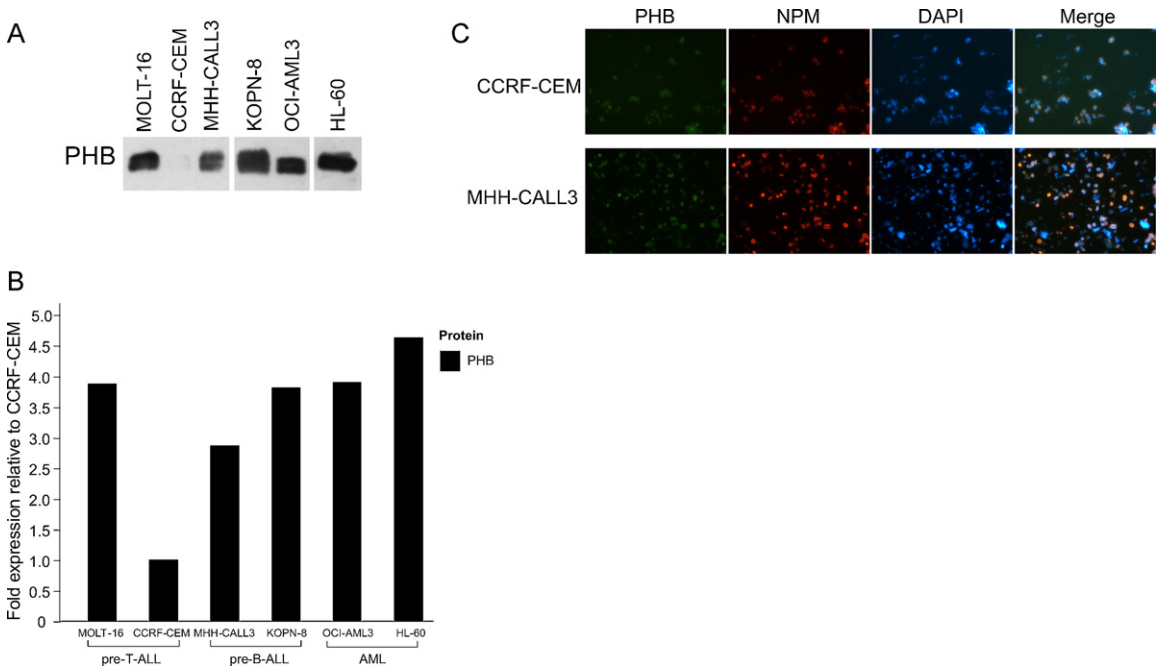

Fig. 2. Nucleolar expression of PHB. (A) Western blot analysis of PHB expression in nucleolar extracts from several leukemic cell lines. Equal loading of samples. (B) Densitometric analysis of the band intensities in panel A normalized to CCRF-CEM levels. (C) Immunofluorescence staining of PHB in nucleolar preparations from CCRF-CEM and MHH-CALL-3 cells. 100× magnification. NPM, nucleophosmin; DAPI, 4′,6-diamidino-2-phenylindole.

expression varies between leukemic cell types. Nucleolar isolations were prepared from various cell lines according to previously published protocols [6]. Nucleophosmin/B23, a well-characterized nucleolar protein [14], was used as a nucleolar marker. As shown in Suppl. Fig. 1, the nucleolar preparations were highly enriched for nucleoli, and only a few sporadic intact nuclei could be seen.

For 2-D DIGE analysis, the experiments were designed to facilitate evaluation of nucleolar protein composition between AML and ALL (OCI-AML3 vs. CCRF-CEM), and between pre-B-ALL and pre-T-ALL (MHH-CALL3 vs. CCRF-CEM). The 2-D DIGE analyses were performed in duplicate. A typical obtained fluorescence image of a 2-D DIGE gel is shown in Fig. 1a. Intensity levels of individual protein spots, mirroring differences in expression levels between various leukemic cell lines, were compared using DeCyder™ software. A number of differentially expressed proteins (average ratio cut-off set to ±2) were detected, 36 in AML/ALL and 29 in pre-B/pre-T cell comparisons. For easier spot picking the 2-DE gels were post-stained with silver (Fig. 1b). Finally, the selected protein spots were "in-gel" digested and the digests subjected to nanoLC–tandem mass spectrometry analysis. Protein identification

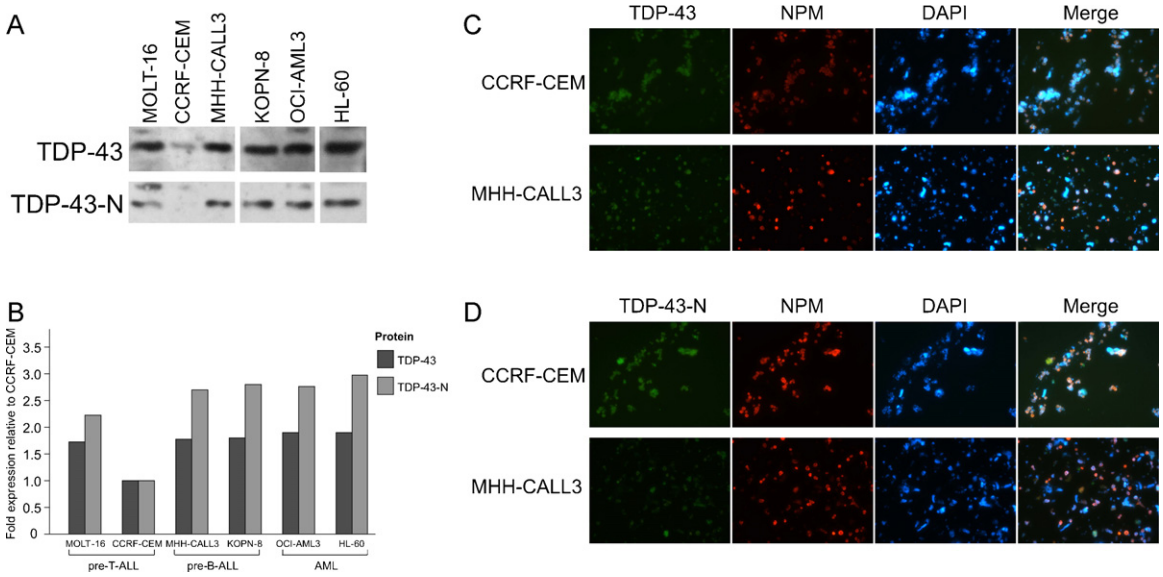

Fig. 3. Nucleolar expression of TDP-43. (A) Western blot analysis of the expression of TDP-43 (TDP-43 and its N-terminal region, TDP-43-N) in nucleolar extracts from leukemic cell lines. Equal loading of samples. (B) Densitometric analysis of the band intensities in panel A normalized to CCRF-CEM levels. (C and D) Immunofluorescence staining of TDP-43 in nucleolar preparations from CCRF-CEM and MHH-CALL3 cells. NPM, nucleophosmin; DAPI, 4′,6-diamidino-2-phenylindole. 100× magnification (C and D).

Table 1
Five most interesting proteins identified with high probability from a pre-B-ALL vs. pre-T-ALL 2-D DIGE gel. Information on the excised spots from DeCyder analyses (av ratio), silver-stained gel (spot type, experimental MW and pI) and mass spectrometric analysis (protein ID and name, MASCOT score and hits). The theoretical MWs and pIs were calculated with ExPASy Compute pI/Mw tool (http://au.expasy.org/tools/pi_tool.html). Prot ID: UniProt accession number, MASCOT score: sum of individual ion scores, peptide matches: the number of peptides identified, MW: molecular weight (kDa), pI: isoelectric point, spot type: silver-staining intensity of the spot on gel, av ratio: average ratio expressing the degree of difference in standardized protein abundance between the 2 cell lines.

spot no	Prot ID	Protein name	MASCOT score	Peptide matches	Experimental		Theoretical		Spot type	Av. ratio
					MW	pI	MW	pI		
2	P56279	T-cell leukemia/lymphoma protein 1A	67	1	14–16	4.8	13.5	5.0	Moderate	9.9
19	O75934	Pre-mRNA-splicing factor SPF27	129	4	28	5.4	26.1	5.5	Moderate	12.3
	P02663	Alpha-S2-casein precursor	54	2	28	5.4	26	8.6	Moderate	12.3
22	P35232	Prohibitin	617	13	30	5.5	29.8	5.6	Strong	6.3
32	Q13148	TAR DNA-binding protein 43	150	2	45	5.8	44.7	5.9	Strong	4.4

was performed using the MASCOT MS/MS Ion Search programme. Four of the protein spots analyzed from pre-B-/pre-T-ALL gels were found to match proteins from the suggested origin with a high probability (Table 1). Two of the proteins, PHB and TDP-43, were deemed the most interesting and we focused further analyses on them.

In order to confirm the 2-D DIGE results, we performed Western blotting of independent nucleolar extracts. As shown in Figs. 2a and 3a, both PHB and TDP-43 were strongly expressed in the nucleolar preparations from a pre-B-ALL cell line (MHH-CALL-3), as compared to a pre-T-ALL cell line (CCRF-CEM). In densitometric analyses, a 1.8–2.7-fold difference in TDP-43 expression and a 2.9-fold difference in PHB expression was seen (Figs. 2b and 3b). This is in line with the 2D-DIGE analyses, in which we observed a 6.3-fold difference in the intensity of the PHB spot and a 4.4-fold difference for the TDP-43 spot. Likewise, another pre-B-ALL cell line (KOPN-8) showed strong expression of both TDP-43 and PHB. On the other hand, another cell line representing pre-T-ALL (MOLT-16) showed stronger expression of TDP-43 and PHB than CCRF-CEM (Figs. 2a and 3a), probably reflecting differences in the genetic background of individual cell lines. Also, two tested AML cell lines (HL-60 and OCI-AML3) demonstrated strong expression of both TDP-43 and PHB (Figs. 2a and 3a).

Finally, we looked at the expression of PHB and TDP-43 more closely by examining their staining patterns by immunofluorescence microscopy. As shown in Fig. 2c, 3c and d, nucleolar extracts of MHH-CALL-3 and CCRF-CEM cells showed strong staining for PHB and TDP-43 (with 2 different antibodies). Both proteins colocalized with nucleophosmin, a nucleolar marker protein.

4. Discussion

This study aimed to identify nucleolar proteins that are differentially expressed in various cell lines derived from acute leukemias. We report here the identification of two such proteins, PHB and TDP-43, which showed strong nucleolar expression in a pre-B-ALL cell line (MHH-CALL-3).

Our methodological approach, which subjects nucleolar extracts to sensitive proteomic analyses, has the ability to capture subtle differences in nucleolar proteomes. Each 2-D DIGE analysis can produce novel findings, and, with the combined internal standard allowing comparisons between gels from separate runs, more extensive analyses are feasible. This is especially convenient in the study of acute leukemia, which comprises multiple disease subtypes. With commercially available cell lines it is possible to identify nucleolar proteins common to all leukemias, or to delineate leukemia subtype-defining expression patterns. A drawback of the strategy used in this study is that large amounts of starting material are needed for the nucleolar extracts, and this limits the application of this method in analysis of patient samples. Nevertheless, the availability of leukemic cell lines makes it possible to perform screening experiments, and the aim is to confirm the obtained results in patient samples.

We identified 2 proteins, PHB and TDP-43, as differentially expressed in the nucleoli of leukemic cell lines. PHB, also known as B-cell receptor associated protein-32 (BAP32), and a closely related protein, PHB2, localize mainly to the mitochondria, where they function as chaperones [15]. However, they have also been found in other subcellular locations, including the plasma membrane and the nucleus, as well as in blood [15–17]. PHB has also been found in the nucleolus (www.lamondlab.com/), consistent with our findings. In the nuclei of breast cancer cells, PHB colocalizes with E2F1, retinoblastoma protein (pRb) and p53, and is able to modulate transcriptional activity mediated by p53 and E2F1 [16–18]. Interestingly, both anti-tumorigenic and pro-tumorigenic functions have been reported for PHB, depending on its subcellular localization [19]. Our results show differences in the nucleolar abundance of PHB between leukemic cell lines. It can be envisioned that the nucleolar localization of PHB is a result of lack of interacting partners in the nucleus (e.g., pRb or p53). PHB may function in regulation of transcription in the nucleolus, in a role similar to its nuclear role. It may also serve as a chaperone in ribosome biogenesis, analogous to its role in mitochondria. Even though PHB also showed strong expression in another pre-B-ALL cell line (KOPN-8), it remains to be seen whether increased nucleolar expression of PHB is common to all pre-B-ALL cells. If this is the case, one may wonder about the reason behind this. On the other hand, the observed differences in expression levels in 2 separate pre-T-ALL cell lines, CCRF-CEM and MOLT-16, could also indicate that other factors may also play a role. It is also noteworthy that the tested AML cell lines showed strong nucleolar expression of both PHB and TDP-43.

TDP-43 belongs to the family of heterogeneous nuclear ribonucleoproteins (hnRNPs) and has been shown to be involved in regulation of pre-mRNA splicing and repression of transcription [20]. Loss of TDP-43 in human cells causes genomic instability and increased apoptosis [21]. Hyperphosphorylated, ubiquitinated, and cleaved TDP-43 was identified as the main component of cytoplasmic inclusions in two neurodegenerative diseases, frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS) [22,23]. The cells containing TDP-43 inclusions show nuclear clearing of TDP-43 [22,24]. Recently, an immunoelectronmicroscopical analysis revealed that TDP-43 localized specifically to the nucleoli in motor neurons [25]. ALS patients had a reduced number of nucleolus-positive deposits, and the authors suggested that TDP-43 trafficking between the nucleus (or nucleolus) and the cytoplasm is disturbed in ALS [25]. In view of our results, disturbed localization of TDP-43 in leukemic cells could link this protein to leukemogenesis. Although TDP-43 involvement in leukemogen-

esis has not been reported previously, our observations suggest that the possible role of TDP-43 in malignancy warrants further studies.

In conclusion, we have demonstrated that the present 2-D DIGE proteomic approach is applicable to the study of leukemia at the nucleolar level. This approach was able to identify proteins that are differentially expressed in different leukemic cell lines. Furthermore, we identified two nucleolar proteins as potential candidates for further studies in the field of leukemia research.

Acknowledgements

We thank Mr. Jorma Kulmala for technical assistance.

Funding source: This work was supported by the Academy of Finland Research Council for Health (funding decision number 115260), the Foundation for Paediatric Research in Finland, the Finnish Medical Foundation, the Competitive Research Funding of Tampere University Hospital (grants 9J062 and 9K073), the Nona and Kullervo Väre Foundation and the Päivikki and Sakari Sohlberg Foundation. The funding sources had no involvement in the study.

Contributions: Kaisa J. Teittinen and Olli Lohi provided conception of the study and drafted the manuscript. Kaisa J. Teittinen, Pauliina Kärkkäinen, Johanna Salonen, Hanna Korkeamäki and Olli Lohi designed the study. Kaisa J. Teittinen, Pauliina Kärkkäinen, Gunilla Rönnholm and Hanna Korkeamäki provided the acquisition of the data. Kaisa J. Teittinen, Pauliina Kärkkäinen, Johanna Salonen, Gunilla Rönnholm, Nisse Kalkkinen and Olli Lohi analyzed the data. Kaisa J. Teittinen, Johanna Salonen, Gunilla Rönnholm, Nisse Kalkkinen and Olli Lohi provided interpretation of the data. Pauliina Kärkkäinen, Johanna Salonen, Gunilla Rönnholm, Hanna Korkeamäki, Mauno Vihinen and Nisse Kalkkinen revised the manuscript and all authors approved the final version of the submitted manuscript.

Conflict of interest: The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.leukres.2011.06.038.

References

[1] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular biology of the cell. 4th ed. NY: Garland Science; 2002.
[2] Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI. The multifunctional nucleolus. Nat Rev Mol Cell Biol 2007;8:574–85.
[3] Lam YW, Trinkle-Mulcahy L, Lamond AI. The nucleolus. J Cell Sci 2005;118:1335–7.
[4] Stark LA, Taliansky M. Old and new faces of the nucleolus. Workshop on the nucleolus and disease. EMBO Rep 2009;10:35–40.
[5] Ahmad Y, Boisvert FM, Gregor P, Cobley A, Lamond AI. NOPdb: nucleolar proteome database—2008 update. Nucleic Acids Res 2009;37(Database issue):D181–4.
[6] Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, et al. Nucleolar proteome dynamics. Nature 2005;433:77–83.
[7] Liu JM, Ellis SR. Ribosomes and marrow failure: coincidental association or molecular paradigm? Blood 2006;107:4583–8.
[8] Grisendi S, Mecucci C, Falini B, Pandolfi PP. Nucleophosmin and cancer. Nat Rev Cancer 2006;6:493–505.
[9] Falini B, Mecucci C, Tiacci E, Alcalay M, Rosati R, Pasqualucci L, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med 2005;352:254–66.
[10] Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, et al. Directed proteomic analysis of the human nucleolus. Curr Biol 2002;12:1–11.
[11] Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH, et al. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis 2000;21:3666–72.
[12] Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. Anal Biochem 1992;203:173–9.
[13] Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, et al. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 1996;379:466–9.
[14] Borer RA, Lehner CF, Eppenberger HM, Nigg EA. Major nucleolar proteins shuttle between nucleus and cytoplasm. Cell 1989;56:379–90.
[15] Mishra S, Murphy LC, Murphy LJ. The prohibitins: emerging roles in diverse functions. J Cell Mol Med 2006;10:353–63.
[16] Wang S, Fusaro G, Padmanabhan J, Chellappan SP. Prohibitin co-localizes with rb in the nucleus and recruits N-CoR and HDAC1 for transcriptional repression. Oncogene 2002;21:8388–96.
[17] Fusaro G, Dasgupta P, Rastogi S, Joshi B, Chellappan S. Prohibitin induces the transcriptional activity of p53 and is exported from the nucleus upon apoptotic signaling. J Biol Chem 2003;278:47853–61.
[18] Wang S, Nath N, Adlam M, Chellappan S. Prohibitin, a potential tumor suppressor, interacts with RB and regulates E2F function. Oncogene 1999;18:3501–10.
[19] Theiss AL, Sitaraman SV. The role and therapeutic potential of prohibitin in disease. Biochim Biophys Acta 2011.
[20] Chen-Plotkin AS, Lee VM, Trojanowski JQ. TAR DNA-binding protein 43 in neurodegenerative disease. Nat Rev Neurol 2010;6:211–20.
[21] Ayala YM, Misteli T, Baralle FE. TDP-43 regulates retinoblastoma protein phosphorylation through the repression of cyclin-dependent kinase 6 expression. Proc Natl Acad Sci U S A 2008;105:3785–9.
[22] Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 2006;314:130–3.
[23] Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun 2006;351:602–11.
[24] Winton MJ, Igaz LM, Wong MM, Kwong LK, Trojanowski JQ, Lee VM. Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. J Biol Chem 2008;283:13302–9.
[25] Sasaki S, Takeda T, Shibata N, Kobayashi M. Alterations in subcellular localization of TDP-43 immunoreactivity in the anterior horns in sporadic amyotrophic lateral sclerosis. Neurosci Lett 2010;478:72–6.

ARTICLE
Journal of Cellular Biochemistry 113:3843–3852 (2012)

Journal of Cellular Biochemistry

SAP30L (Sin3A-Associated Protein 30-Like) is Involved in Regulation of Cardiac Development and Hematopoiesis in Zebrafish Embryos

Kaisa J. Teittinen,[1*] Toni Grönroos,[1] Mataleena Parikka,[2] Sini Junttila,[3] Annemari Uusimäki,[1,2] Asta Laiho,[3] Hanna Korkeamäki,[1] Kalle Kurppa,[1,4] Hannu Turpeinen,[2] Marko Pesu,[2,5] Attila Gyenesei,[3] Mika Rämet,[2,4] and Olli Lohi[1,4]

[1]*Paediatric Research Centre, University of Tampere School of Medicine and Tampere University Hospital, 33520 Tampere, Finland*

[2]*Institute of Biomedical Technology, BioMediTech, University of Tampere, 33520 Tampere, Finland*

[3]*Turku Centre for Biotechnology and the Finnish Microarray and Sequencing Centre, University of Turku and Åbo Akademi University, 20520 Turku, Finland*

[4]*Department of Paediatrics, Tampere University Hospital, 33521 Tampere, Finland*

[5]*Fimlab Laboratories, Tampere University Hospital, 33521 Tampere, Finland*

ABSTRACT

The Sin3A-associated proteins SAP30 and SAP30L share 70% sequence identity and are part of the multiprotein Sin3A corepressor complex. They participate in gene repression events by linking members of the complex and stabilizing interactions among the protein members as well as between proteins and DNA. While most organisms have both SAP30 and SAP30L, the zebrafish is exceptional because it only has SAP30L. Here we demonstrate that SAP30L is expressed ubiquitously in embryonic and adult zebrafish tissues. Knockdown of SAP30L using morpholino-mediated technology resulted in a morphant phenotype manifesting as cardiac insufficiency and defective hemoglobinization of red blood cells. A microarray analysis of gene expression in SAP30L morphant embryos revealed changes in the expression of genes involved in regulation of transcription, TGF-beta signaling, Wnt-family transcription factors, and nuclear genes encoding mitochondrial proteins. The expression of the heart-specific nkx2.5 gene was markedly down-regulated in SAP30L morphants, and the cardiac phenotype could be partially rescued by nkx2.5 mRNA. In addition, changes were detected in the expression of genes known to be important in hemoglobin synthesis and erythropoiesis. Our results demonstrate that SAP30L regulates several transcriptional pathways in zebrafish embryos and is involved in the development of cardiac and hematopoietic systems. J. Cell. Biochem. 113: 3843–3852, 2012.

KEY WORDS: ZEBRAFISH; SAP30L; TRANSCRIPTION; MICROARRAY; HEMATOPOIESIS; CARDIOGENESIS

Acetylation and other covalent modifications of histones play fundamental roles in chromatin dynamics and regulation of gene expression. Deacetylation of histones is carried out by a multiprotein corepressor complex in which Sin3A is an essential scaffold protein. Sin3A is composed of domains that mediate protein–protein interactions and thereby forms a platform to which several enzymes, DNA-binding transcription factors, and other bridging proteins bind [Silverstein and Ekwall, 2005]. The core

The authors have no conflict of interest to declare.

Additional supporting information may be found in the online version of this article.

Grant sponsor: Academy of Finland Research Council for Health; Grant numbers: 115260, 121003, 1286223, 135980, 135736, 139225; Grant sponsor: Foundation for Paediatric Research in Finland; Grant sponsor: Finnish Medical Foundation; Grant sponsor: Competitive Research Funding of Tampere University Hospital; Grant numbers: 9J062, 9K073, 9M052, 9L070, 9K093, 9L075, 9M080, 9M093; Grant sponsor: Marie Curie International Reintegration Grant within the 7th European Community Framework Programme; Grant sponsor: Emil Aaltonen Foundation; Grant sponsor: Sigrid Juselius Foundation; Grant sponsor: Nona and Kullervo Väre Foundation; Grant sponsor: Päivikki and Sakari Sohlberg Foundation.

*Correspondence to: Kaisa J. Teittinen, Paediatric Research Centre, University of Tampere School of Medicine, FIN-33520 Tampere, Finland E-mail: kaisa.teittinen@uta.fi

Manuscript Received: 28 May 2012; Manuscript Accepted: 16 July 2012

Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 20 July 2012

DOI 10.1002/jcb.24298 • © 2012 Wiley Periodicals, Inc.

Sin3A-HDAC corepressor complex contains the histone deacetylases HDAC1 and HDAC2, the histone-binding proteins RbAp46 and RbAp48, the Sin3A-associated protein 18 (SAP18), SAP30, and SDS3 [Silverstein and Ekwall, 2005].

The Sin3A-associated proteins SAP30 and SAP30L (SAP30-like) have been shown to be present in various corepressor complexes [Laherty et al., 1998; Zhang et al., 1998; Viiri et al., 2006]. We have previously reported that these proteins bind DNA directly and this association is regulated by nuclear monophospholipids [Viiri et al., 2009b]. SAP30 and SAP30L are small and basic, and contain nuclear (NLS) and nucleolar (NoLS) localization signals [Viiri et al., 2006]. The main difference between the two proteins lies in their N-termini, where SAP30 contains a 38 amino acid insertion relative to SAP30L. Both proteins are able to localize to the nucleus or the nucleolus, and they can direct Sin3A to the nucleolus [Viiri et al., 2006].

While a range of biochemical and cell biological results implicate SAP30 and SAP30L in repression of gene expression, the function(s) of these proteins in the setting of a living organism has not been established. In order to study this, we chose the zebrafish as a model organism. The zebrafish has only one of these proteins, SAP30L [Viiri et al., 2009a], enabling us to study the role of SAP30L without interference of potentially overlapping functions of SAP30. Our results suggest an important role for SAP30L in diverse transcriptional pathways in zebrafish embryos.

MATERIALS AND METHODS

ZEBRAFISH MAINTENANCE AND MICROINJECTIONS

Wild type AB zebrafish were maintained under standard conditions at 28.5°C as described [Westerfield, 1995]. In microinjection experiments, embryos at the 1-4-cell stage were injected into the yolk sac with 1 nl of injection mix, which contained the morpholino, mRNA (where indicated) and rhodamine dextran. The PV830 Pneumatic PicoPump injection system (WPI, Inc.) was used, and the injection volume was calibrated to 1 nl using 1 mm Micrometer scale (S48, Ted Pella Inc.). The care of the animals was in accordance with the Finnish Laboratory Animal Welfare Act 62/2006 and the Laboratory Animal Welfare Ordinance 36/2006.

WHOLE-MOUNT IN SITU HYBRIDIZATION (WISH)

DIG-labeled sense and antisense probes for zebrafish SAP30L were prepared using the DIG RNA Labeling Mix (Roche Diagnostics) according to the manufacturer's instructions, as described [Thisse and Thisse, 2008]. Linearized plasmids containing the full-length SAP30L cDNA (in the sense or antisense orientation) under the control of the T7 promoter were used as templates. Zebrafish larvae were collected and processed as described [Thisse and Thisse, 2008] at 12, 24, 36, 48, and 72 hpf. Whole-mount in situ hybridization (WISH) using alkaline phosphatase detection with the BM Purple substrate (Roche Diagnostics) was carried out according to Thisse and Thisse [2008] with the following modifications: 2 ml tubes were used during the entire procedure, hybridization was carried out at +65°C, 130 ng of sense or antisense probe was used per reaction and a dilution ratio of 1:2,000 was used for the anti-DIG antibody (sheep anti-digoxigenin-AP Fab fragments, Roche Diagnostics). The stained fish were analyzed and photographed under a dissecting microscope. The localization of the staining was assessed by comparing to the anatomical reference images by Haffter et al. [1996].

RNA EXTRACTION AND QUANTITATIVE REAL-TIME PCR

Expression of the SAP30L, nkx2.5 and ppox genes was assayed by quantitative real-time PCR (qRT-PCR). Total RNA was extracted from zebrafish embryos and adult fish tissues with an RNeasy Mini Kit (Qiagen), and converted to cDNA using an iScript Select cDNA Synthesis Kit (Bio-Rad) and random primers according to the manufacturers' instructions. The housekeeping gene EF1A (Ensembl ID: ENSDARG00000020850) [Tang et al., 2007] was used as a normalization control. The primer sequences are described in the Supplementary data. qRT-PCR was performed by using an Evagreen Ssofast supermix kit (Bio-Rad) as instructed by the manufacturer. An average threshold cycle (Ct) value was calculated from two or three replicate samples. The mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. The PCR runs were repeated twice, and no significant differences between replicate runs were observed (data not shown). For the microarray experiments, the total RNA was extracted from whole embryos collected at 12 or 24 hpf.

ANTISENSE MORPHOLINO EXPERIMENTS

The translational start site of the zebrafish SAP30L mRNA (Ensembl ID: ENSDARG00000030213) was targeted with two independent antisense morpholinos: SAP30L-MO1 and SAP30L-MO2 (Fig. 2a, Supplementary Table I). Their sequences exhibited no significant similarity to other loci in a search of a zebrafish database. The random control morpholino (RC, see Supplementary Table I) has no gene targets or observable biological activity in the zebrafish. The p53-targeting morpholino was also used (Supplementary Table I). All morpholinos were purchased from Gene-Tools, LLC. The effective concentrations of morpholinos were experimentally titrated and are indicated in the figures or figure legends.

WESTERN BLOTTING

The efficacy of MO1 and MO2 in blocking the translation of the SAP30L mRNA was analyzed using a human anti-SAP30L antibody [Korkeamaki et al., 2008]. MO1-, MO2-, or RC-injected larvae were collected at 2 dpf (days post-fertilization) directly into 2× Laemmli buffer (10 larvae/50 μl), and lysed by boiling and passing through a pipette vigorously. The lysates were centrifuged for 5 min to remove the debris, and the soluble fraction was used for further analysis. The proteins were resolved on 12% SDS–PAGE gels and transferred to nitrocellulose membranes (Amersham Biosciences). The immunoblots were probed with an anti-SAP30L antibody, and the proteins were detected by enhanced chemiluminescence. In all zebrafish lysates, the anti-SAP30L-antibody recognized only a single band, the molecular weight of which is a few kilodaltons lower than in human control (lysed cells transfected with Myc-His-tagged SAP30L), as expected due to the presence of the tag and the larger size of human SAP30L.

ANALYSIS OF HEART MORPHOLOGY AND FUNCTION

Heart morphology at 5 dpf was assessed visually and categorized as normal or deformed, examples of which are shown in Supplementary Figure S3. Heart rates were determined visually in unanaesthetized morphant and control larvae at 5 dpf. Ventricular and atrial performance was analyzed on 10–20 s time-lapse videos. The longitudinal fractional shortening was calculated according to the formula $(LDd - LDs)/LDd$, where *LDd* is the longitudinal diameter at diastole and *LDs* the diameter at systole.

HISTOLOGICAL STAINING

The morphant and control larvae were fixed in a 4% formaldehyde–PBS solution at +4°C and embedded in 2% agarose. The samples were dehydrated by incubating them in a series of alcohol solutions (70%, 96%, and absolute ethanol) for 1–2 h in each solution, and finally in xylene for at least 1 h. The samples were then embedded in paraffin and 5 μm-thick sections were cut in longitudinal orientation. The sections were fixed on glass slides and deparaffinised by incubating them twice in xylene (á 4 min), twice in absolute ethanol (á 3 min), twice in 96% ethanol (á 3 min), once in 70% ethanol, and rinsing once with water. Hematoxylin–eosin staining was performed by incubating the samples in Mayer's hematoxylin (2 min), running water (2 min), water (1.5 min), 70% ethanol (15 s), eosin Y (15 s), 96% ethanol (30 s, twice), absolute ethanol (1 min, twice), and xylene (1–4 min). After this the slides were mounted and pictures were taken under a BX60 microscope using the CellˆD program (Olympus).

O-DIANISIDINE STAINING

Detection of hemoglobin by *o*-dianisidine staining was performed as described previously [Ransom et al., 1996]. Larvae stained at 3 dpf were scored as normal, reduced, or severely reduced based on their qualitative level of staining which was assessed visually. In the normal category, staining was judged to be equivalent to that of wild type larvae, reduced staining was visibly decreased as compared to wild type, and in the severely reduced category, the larvae showed little or no staining.

mRNA RESCUE

The mRNA rescue experiments were performed by injecting simultaneously SAP30L-MO1 or SAP30L-MO2 and a capped synthetic mRNA encoding the gene of interest. We used various doses of mRNA in the injections, and the data were pooled for analysis. The mRNA was synthesized using a T7 mMESSAGE mMACHINE Kit (Ambion) according to the manufacturer's instructions. Briefly, a plasmid template linearized with *Hind*III was mixed thoroughly with components of the kit and incubated at +37°C for 2 h, after which the template DNA was removed by a DNase treatment. The synthesized mRNA was purified using a MEGAclear kit (Ambion) according to the manufacturer's protocol and further concentrated by ammonium acetate precipitation. The templates used in the synthesis reactions had been created by cloning the desired cDNA into the pcDNA3.1/*myc*His(−)A-vector (Invitrogen) using *Eco*RI and *Hind*III restriction sites. The constructs contained full-length zebrafish SAP30L (drSAP30L) or full-length zebrafish nkx2.5 (Nkx2.5), each without a stop codon.

STATISTICAL ANALYSES

Quantitative data are given as mean ± standard deviation. When appropriate, differences between groups were compared by using the two-tailed Student *t*-test or the χ^2 test for cross tabulations. A *P*-value of less than 0.05 was considered statistically significant. Statistical calculations were performed using PASW Statistics version 18.

MICROARRAYS

200 ng of total RNA was amplified and labeled with Cy3 using a Low Input Quick Amp Labeling kit (Agilent Technologies). The samples were processed using an RNA Spike In kit (Agilent Technologies). 1.65 μg of each Cy3-labeled sample was hybridized to Agilent's 4 × 44K Zebrafish V3 Gene Expression Microarray overnight at 65°C in the buffers of Agilent's Gene Expression Hybridization kit. The arrays were washed according to the manufacturer's instructions and scanned with an Agilent Technologies' scanner (model G2565CA), using scan profile AgilentHD_GX_1Color. Numerical data were extracted with Agilent's Feature Extraction software, version 10.7.1., using grid 026437_D_F_20100719 and protocol GE1_107_Sep09.

MICROARRAY DATA ANALYSIS

The R programming language and environment [R Development Core Team, 2008] and its Bioconductor module [Gentleman et al., 2004] were used for analysis of data from the microarrays. The raw data were normalized with quantile normalization, and the quality of the data was checked with several quality control methods. Statistical testing was performed with the Limma package [Smyth, 2005], and stringent filtering thresholds were used for both time point comparisons (fold change FC$\geq$|3| and *P*-value $P \leq 0.001$).

RESULTS AND DISCUSSION

SAP30L mRNA IS EXPRESSED DURING EMBRYOGENESIS AND IN ADULT ZEBRAFISH TISSUES

SAP30 and SAP30L have been implicated in the regulation of gene repression through the Sin3A-corepressor complex [Laherty et al., 1998; Zhang et al., 1998; Viiri et al., 2006, 2009b]. Their function in vivo remains unknown and we therefore set out to examine the function of SAP30L using the zebrafish as a model. Contrary to the human and mouse genomes, only one member of the SAP30 family has thus far been identified in the zebrafish genome, and its derived protein sequence more closely resembles human SAP30L than SAP30 (Supplementary Fig. S1; Viiri et al., 2009a).

We studied the expression of the SAP30L gene during zebrafish embryogenesis in whole embryos by qRT-PCR, and determined the localization of the expression by WISH at 12, 24, 36, 48, and 72 h post-fertilization (hpf) (Fig. 1a and b). The results show that SAP30L mRNA is widely expressed during embryogenesis (Fig. 1a and b). The expression is strongest in the brain, including forebrain, midbrain, and hindbrain, at the 24, 36, 48, and 72 hpf time points (Fig. 1a). The pectoral fin buds (at 36 and 48 hpf) and pectoral fins (72 hpf) also show strong expression of the mRNA. Next to the pectoral fin buds, staining is seen also in the region of common cardinal vein (see dorsal view in Fig. 1a) starting at 36 hpf. Furthermore, the heart area

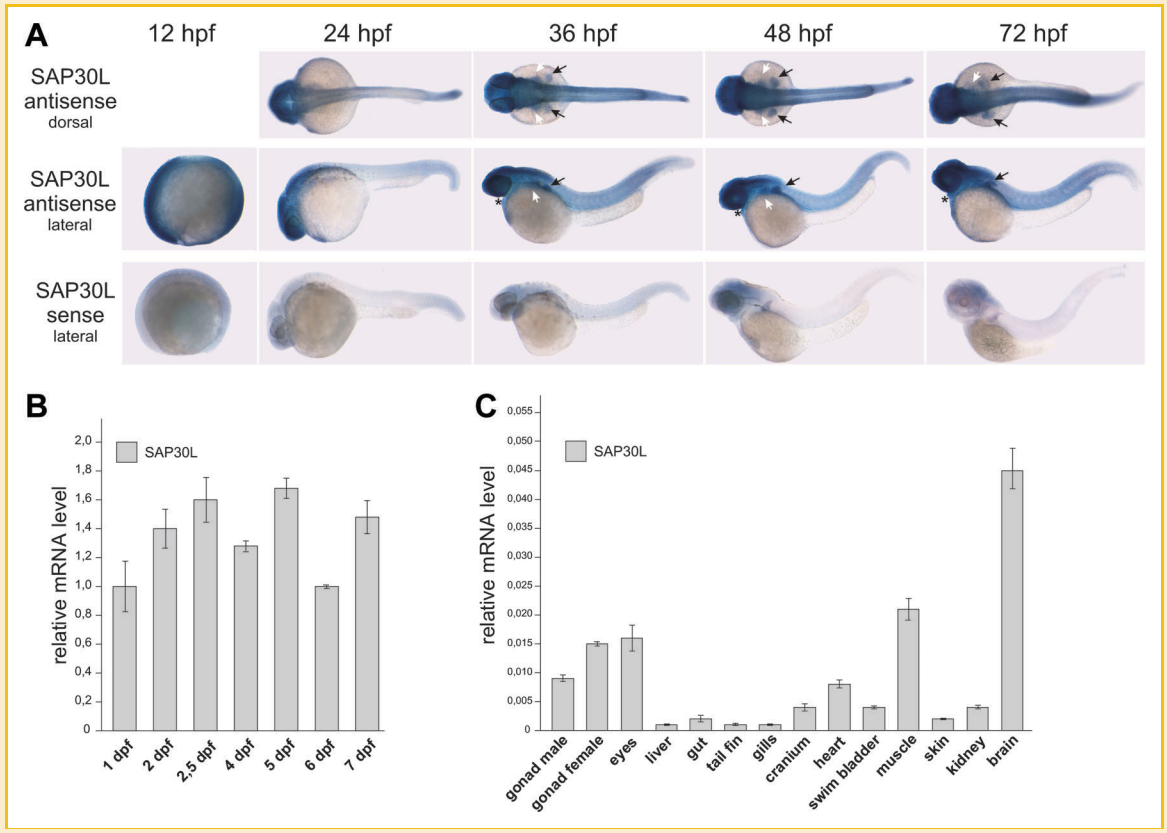

Fig. 1. Expression of SAP30L mRNA in zebrafish embryos, larvae, and adult tissues. A: WISH analysis of SAP30L mRNA expression during embryogenesis at indicated time points. Staining with an antisense probe is shown in dorsal (top row) and lateral (middle row) views, and a control sense-strand probe is shown in lateral view (bottom row). The black asterisk points out the heart next to it, the black arrow indicates the pectoral fin bud/pectoral fin, and the white arrow indicates the region of the common cardinal vein. B: qRT-PCR analysis of SAP30L mRNA in whole embryos during embryogenesis. SAP30L mRNA levels are normalized to the level at 1 day post-fertilization (dpf). C: qRT-PCR analysis of SAP30L mRNA in adult tissues. SAP30L mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

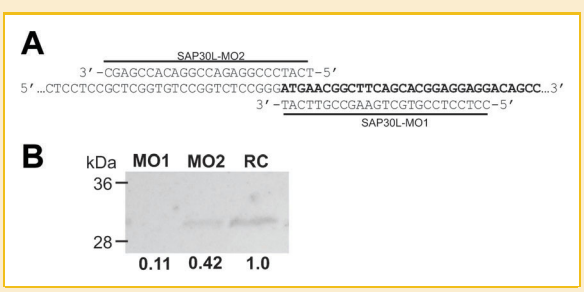

Fig. 2. The morpholinos used to target the SAP30L mRNA. A: Sequence of the coding strand of the zebrafish SAP30L gene, showing the 5′-untranslated region (normal font) and part of the first exon (boldface starting from the ATG-codon). Complementarity of the sequences of the morpholinos SAP30L-MO1 and SAP30L-MO2 to the coding strand of the gene is illustrated. B: Western blotting analysis of SAP30L protein levels in 2 dpf larvae injected with SAP30L-MO1 (MO1), SAP30L-MO2 (MO2), or random control morpholino (RC). Morpholino concentrations used: 400 μM for MO1, 600 μM for MO2, 600 μM for RC. The numbers under the lanes indicate band intensity relative to RC. The lysates contain identical numbers of larvae per volume of lysis buffer, and equal volume of sample was loaded on the gel in each case.

shows positive staining, first seen in the pericardial cavity (24 hpf) and later in the heart itself. In addition, moderate expression of SAP30L mRNA is seen in the trunk (including the area of the intermediate cell mass, ICM), and some embryos show weak staining in the blood island at 24 hpf. In adult fish, strong expression is seen in the brain and moderate expression in muscle, eye, gonad, and heart tissues (Fig. 1c).

DESIGN AND VALIDATION OF THE SAP30L-TARGETING MORPHOLINOS

In zebrafish embryos, morpholino oligonucleotides (MO) can be used to knock down expression of genes of interest. We designed two translation-blocking morpholinos (MO1 and MO2) which target the translation initiation site in the SAP30L mRNA (Fig. 2a). In an in vitro translation assay, a specific and dose-dependent inhibition of translation of SAP30L mRNA was detected with the MO1 but not with RC control (Supplementary Fig. S2). The ability of morpholinos to block translation of SAP30L was further investigated by using the SAP30L antibody and Western blotting. A significant reduction in SAP30L protein levels was detected in MO1- and MO2-treated zebrafish embryos at 2 dpf, demonstrating the efficacy of these

morpholinos in down-regulating the expression of the SAP30L protein in vivo (Fig. 2b).

KNOCKDOWN OF SAP30L EVOKES CARDIAC DEFECTS AND REDUCED HEMOGLOBIN LEVELS

In order to investigate the role of SAP30L in zebrafish development, we injected morpholinos (MO1 or MO2) into 1- to 4-cell stage embryos, and observed their development for 5–6 dpf. The SAP30L morphants showed prominent and progressive pericardial edema. Importantly, most of the SAP30L morphants exhibited deformed cardiac morphology (85–89%; Fig. 3a, Supplementary Fig. S3), with the hearts remaining string-like and visible on a single plane, in contrast to the three-dimensional complexity of the healthy hearts in the control larvae (Fig. 3b). Cardiac function was assessed by determining the heart rates and the contractility of the chambers. The results demonstrate significantly reduced heart rates in MO1- and MO2-injected morphants (Fig. 3c), with the differences detectable already at 3 dpf. Similarly, a substantial decrease in the longitudinal fractional shortening values was seen in both the atrium and the ventricle, as compared to control hearts (Fig. 3d). Videos illustrating the diminished cardiac function can be found in the Supplementary data (SV1-3). The results are compatible with the WISH and qRT-PCR analyses which showed expression of SAP30L in the developing and adult heart (Fig. 1a and c), and suggest that the phenotype results specifically from the knockdown of SAP30L.

In some SAP30L morphants, the red blood cells seemed paler as compared to those in control larvae. In a Sin3B-knockout mouse, impaired erythroid differentiation has been reported, with a 43% reduction in the number of circulating red blood cells and a 37% reduction in hemoglobin levels [David et al., 2008]. In order to determine hemoglobin levels in the SAP30L morphants, *o*-dianisidine staining was performed on 3 dpf embryos (see examples in Supplementary Fig. S4). As summarized in Figure 3e, the morphants exhibited reduced staining, indicating reduced hemoglobin levels in the red blood cells. The observed effect on erythroid differentiation is consistent with the fact that SAP30L expression is seen at the area of the common cardinal vein and ICM,

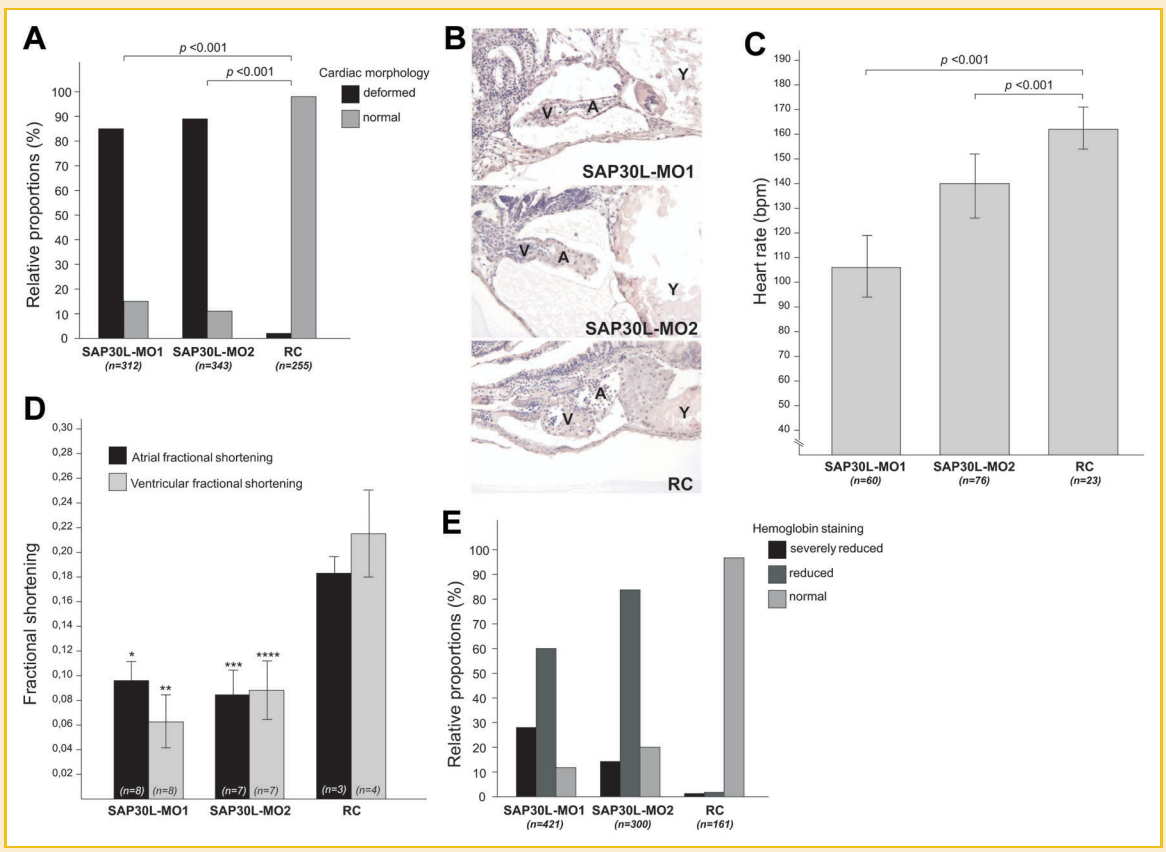

Fig. 3. The phenotype induced by knockdown of SAP30L involves defects in the heart and in hemoglobinization of red blood cells. A: The majority of the SAP30L morphant larvae have deformed hearts (see close-up images of normal and deformed hearts in Supplementary Fig. S2). B: Histological analysis of cardiac morphology. Hematoxylin–eosin staining of sections reveals string-like appearance of the heart in the MO1- and MO2-injected larvae, in contrast to the RC-injected controls. V, ventricle; A, atrium; Y, yolk sac. C: SAP30L morphant larvae exhibit reduced heart rates. Heart rates (beats per minute, bpm) are represented as mean bpm ± SD. D: Both atrial and ventricular performance in SAP30L morphant hearts are impaired, as shown by the decreased longitudinal fractional shortening values (represented as mean ± SD). $^{*}P = 0.002$; $^{**}P < 0.001$; $^{***}P = 0.005$; $^{****}P = 0.006$, as compared to RC control. The morpholino concentrations used in A–D: 300–750 μM for MO1, 500–600 μM for MO2, 400–750 μM for RC. E: SAP30L morphants exhibit defects in hemoglobinization of red blood cells (see examples in Supplementary Fig. S4). MO1-, MO2-, or RC-injected larvae are scored as normal, reduced, or severely reduced based on hemoglobin staining by *o*-dianisidine. Morpholino-concentrations used: 400 μM for MO1, 600 μM for MO2, 600 μM for RC. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

where developing erythrocytes are present [Davidson and Zon, 2004].

In addition, some morphant larvae were slightly smaller than control larvae, and showed edema in the head and brain (data not shown).

The phenotypes obtained with both SAP30L-targeting morpholinos were essentially similar, with MO2-morphants showing slightly delayed onset, which concurs with the less decreased level of SAP30L protein in these morphants as compared to MO1-morphants (Fig. 2b). Since morpholinos are known to induce off-target p53-mediated effects [Robu et al., 2007], we co-injected both morpholinos with the p53-morpholino. Larvae injected with SAP30L-MO2 alone or together with the p53-morpholino, did not differ in phenotype. SAP30L-MO1 morphants exhibited a slightly alleviated but essentially similar phenotype when co-injected with the p53-morpholino (data not shown).

One way to demonstrate the specificity of the observed phenotype is to rescue it by co-injecting the morpholino with the target mRNA [Eisen and Smith, 2008]. In our case, this strategy could be exploited only with MO2, since it targets the SAP30L mRNA upstream of the start codon. No rescue of the reduced hemoglobin staining but a tendency towards a less severe cardiac phenotype could be observed in the rescued morphant larvae (data not shown). The incomplete rescue may be due to various reasons, such as the short half-life of the injected mRNA, strict temporal or spatial constraints on the expression at certain developmental stages, inability of the mRNA to reach the right tissue, or the fine balance between under- and over-expression of the protein in the embryos and larvae.

SAP30L IS INVOLVED IN MULTIPLE TRANSCRIPTIONAL PATHWAYS DURING ZEBRAFISH EMBRYOGENESIS

SAP30L is a part of the Sin3A-HDAC multiprotein corepressor complex [Viiri et al., 2006]. In order to elucidate the transcriptional pathways in which SAP30L participates during zebrafish embryogenesis, we performed a microarray analysis of embryos in which expression of SAP30L was knocked down. To avoid any secondary effects on gene expression, we chose early time points at 12 and 24 hpf. It has been reported previously that most developmentally regulated genes reach their peak expression by 24 hpf, with a sharp increase in transcript accumulation at 12 hpf [Mathavan et al., 2005]. In our analysis, 205 genes were up-regulated and 174 genes down-regulated at the 12 hpf time point, whereas at 24 hpf, the respective numbers were 143 and 64 (Fig. 4, Supplementary Fig. S5). Of the affected genes, 46 were identified at both time points and showed similar alterations in expression. Since SAP30L is part of a transcriptional repressor complex, its knockdown may be expected to lead to more up-regulation than down-regulation of genes. Our results bear out this expectation and are in line with the published microarray data for Sin3, another member of the same repressor complex. RNAi-mediated knockdown of *Drosophila* SIN3 resulted in induction of 364 genes but only 35 genes were repressed [Pile et al., 2003]. Similarly, in mouse cells in which Sin3A had been knocked out, the majority of the differentially expressed genes were up-regulated [Dannenberg et al., 2005]. The results of knockdown of *Drosophila* SIN and the mouse mSin3A were similar in that mainly genes involved in cytosolic and mitochondrial energy-generating pathways were affected. Our data on knockdown of SAP30L in *Danio rerio* show no marked changes in expression of the genes reported in the *Drosophila* SIN and mouse Sin3A data sets. Instead, expression of several nuclear genes encoding mitochondrial proteins was affected by knockdown of SAP30L. These include timm8a, cyb5b, ppox, alas2 and fars2 (Fig. 4, Supplementary Fig. S5).

SAP30L was originally identified in our laboratory as a TGF-beta-inducible transcript in a cell culture system that models differentiation of the intestinal epithelium [Lindfors et al., 2003]. Interestingly, down-regulation of SAP30L in the zebrafish resulted in strong induction of Smad2, a mediator in the TGF-beta signaling pathway. The 150 kDa TGF-beta-1-binding protein was similarly up-regulated at 12 hpf, whereas the bone morphogenetic factor 5 (BMP5), a member of the TGF-beta superfamily, was slightly down-regulated. These results taken together with our previous findings suggest a role for SAP30L in TGF-beta signaling events.

In gene ontology analysis, the pathway with the most hits is associated with regulation of transcription. The same result was obtained when the analysis was performed on the genes that showed the largest differences in their expression. Indeed, several transcription factors show altered expression in our microarray analysis (Fig. 4, Supplementary Fig. S5). Among them are the Wnt-family transcription factor Wnt7b, which is involved in the development of vasculature, lung, and kidney [Shu et al., 2002; Lobov et al., 2005; Rajagopal et al., 2008; Yu et al., 2009; Lin et al., 2010], and Wnt2, which has been shown to be associated with cardiogenesis and several other developmental processes [Alexandrovich et al., 2006; Wang et al., 2007; Goss et al., 2009, 2011; Sousa et al., 2010]. In addition, Wisp3, a protein involved in modulation of Wnt and BMP signaling [Nakamura et al., 2007], was affected.

NKX2.5 EXPRESSION IS REDUCED IN SAP30L KNOCKDOWN EMBRYOS

One of the genes down-regulated at 12 hpf in the microarray analysis of SAP30L morphants, Nkx2.5 (Fig. 4), is a well-known regulator of heart development [Chen and Fishman, 1996; Targoff et al., 2008; Reamon-Buettner and Borlak, 2010]. The down-regulation of the Nkx2.5 gene was confirmed by qRT-PCR, which showed a threefold decrease in its expression in SAP30L-MO1-treated embryos, and a sixfold decrease in SAP30L-MO2-treated embryos (Fig. 5a). Interestingly, co-injection of Nkx2.5 mRNA with SAP30L-MO1 could partially rescue the deformed cardiac morphology in the SAP30L morphants (Fig. 5b). Previous studies in other model organisms have implicated Nkx2.5 in various cardiac developmental processes, including progenitor specification and proliferation, heart tube extension and looping, and chamber morphogenesis [Targoff et al., 2008]. Our results suggest that transcription of Nkx2.5 is regulated by a SAP30L-containing repressor complex in zebrafish embryos. As SAP30L is a member of a large corepressor complex and binds DNA only nonspecifically, its observed effect on expression of Nkx2.5 is likely to be indirect. Interestingly, some members of the Sin3A corepressor complex are known to regulate expression of genes critical for cardiac morphogenesis and function in both the zebrafish and mice. These include HDAC1 [Pillai et al., 2004; Montgomery et al., 2007], HDAC2

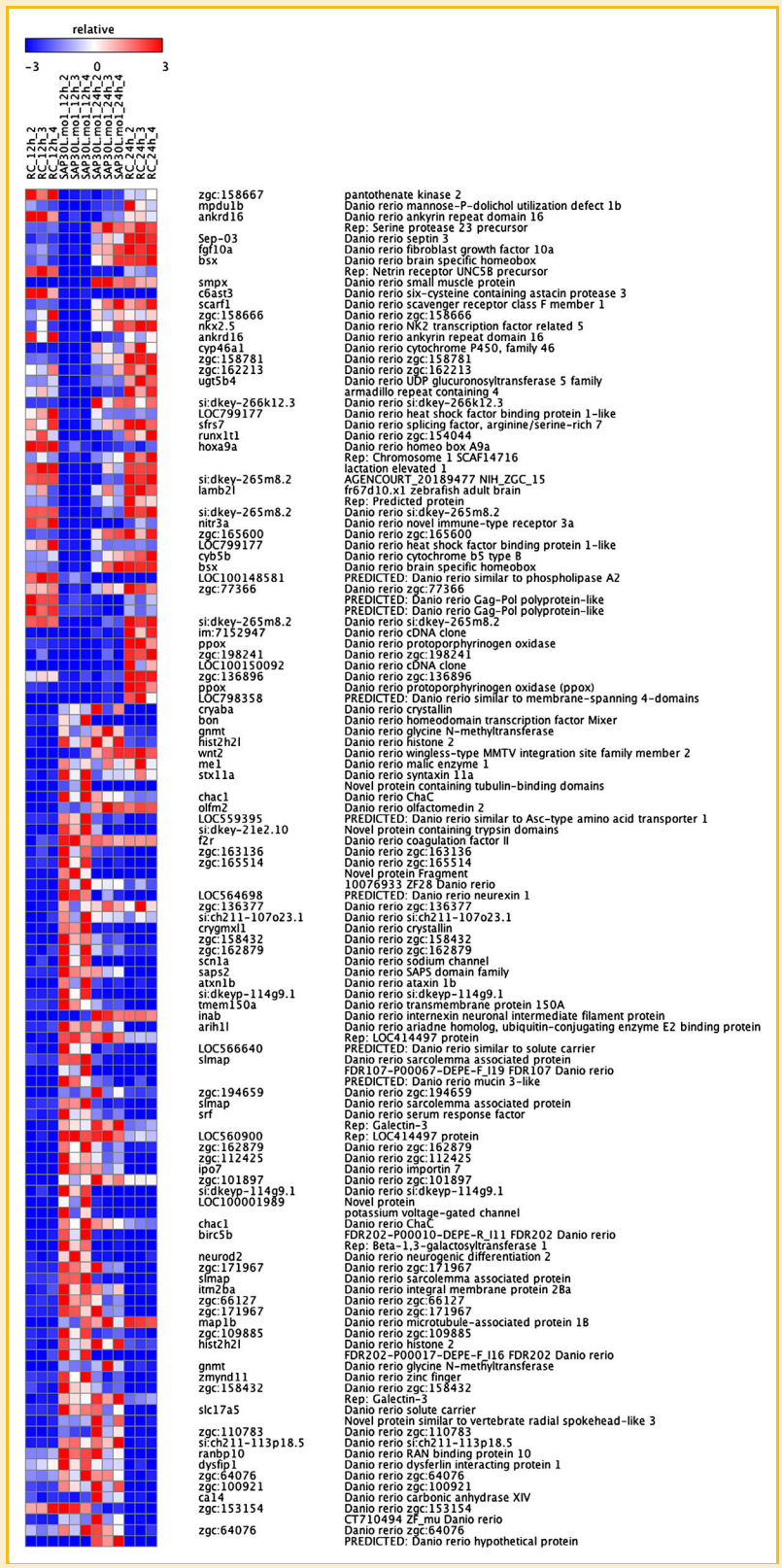

Fig. 4. Microarray analyses of zebrafish embryos (at 12 and 24 hpf) treated with SAP30L-MO1 or random control morpholino (RC). The 104 genes showing the largest differences in expression, ranked according to both the fold-change value and the *P*-value are shown for both time points and in all three parallel samples (2, 3, 4). Blue color denotes down-regulation, red up-regulation, and white no change. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

[Montgomery et al., 2007], HDAC3 [Farooq et al., 2008], and mSDS [Amsterdam et al., 2004].

GENES INVOLVED HEMOGLOBIN SYNTHESIS ARE DOWN-REGULATED BY SAP30L KNOCKDOWN

Consistent with the finding of diminished hemoglobin staining in the morphant larvae, we observed a marked reduction in expression of genes involved in hemoglobin synthesis and erythropoiesis (ppox, alas2, and ba1l globin-like gene) at 24 hpf (Fig. 4, Supplementary Fig. S5). The zebrafish ppox, a protoporphyrinogen oxidase, is expressed in the ICM as early as 22 hpf, similar to other genes involved in hematopoiesis, including alas2, gata1, and β-spectrin [Detrich et al., 1995; Brownlie et al., 1998; Liao et al., 2000]. Down-regulation of ppox was confirmed by qRT-PCR, which showed a sixfold decrease in its expression in SAP30L-MO1-treated embryos and a 4.8-fold decrease in SAP30L-MO2-treated embryos (Fig. 5c). In the future, it remains to be elucidated whether the lower hemoglobin levels observed after knockdown of SAP30L are due to impaired erythropoiesis, a block in heme biosynthesis, or some other mechanism. Our results suggest that one function of SAP30L may be to silence the expression of a factor that inhibits heme synthesis or erythropoiesis.

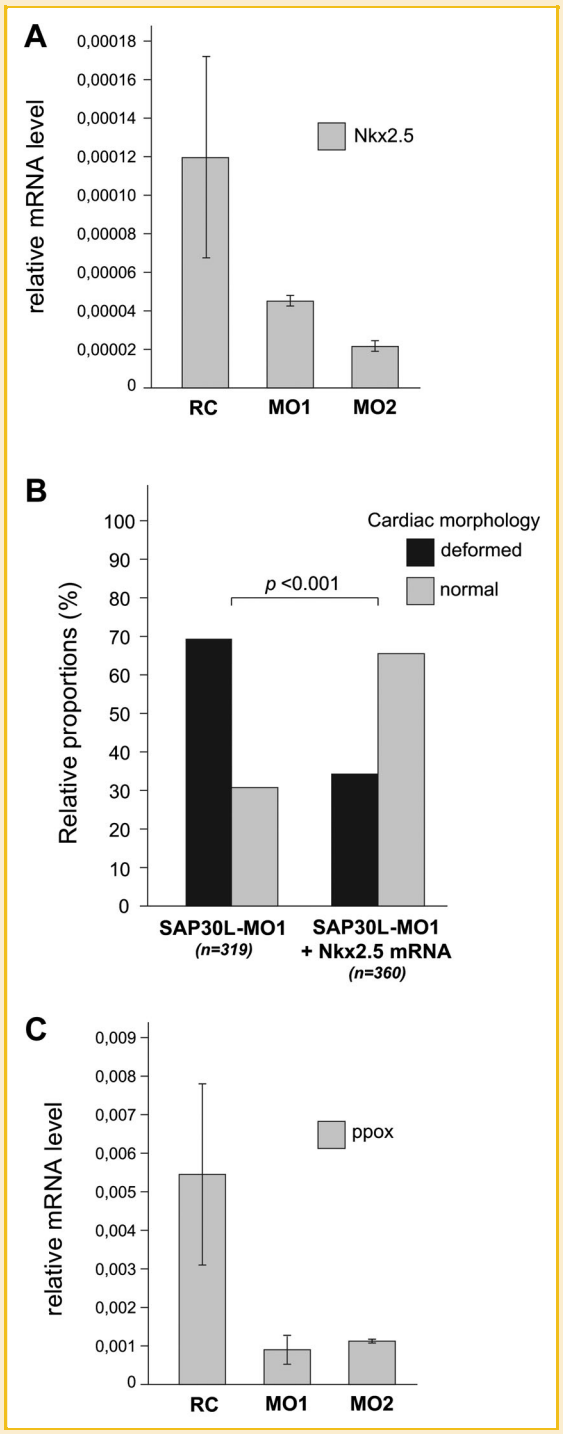

Fig. 5. Expression of Nkx2.5 and ppox mRNAs is down-regulated during SAP30L knockdown. A: Nkx2.5 mRNA expression is reduced in both SAP30L-MO1- and SAP30L-MO2-treated embryos at 12 hpf relative to RC-treated embryos. Nkx2.5 mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. B: The number of larvae exhibiting cardiac deformation is significantly reduced when treated with both SAP30L-MO1 (400 μM) and Nkx2.5 mRNA (20–50 pg per embryo), as compared to larvae treated with SAP30L-MO1 only. C: ppox mRNA expression is reduced in SAP30L-MO1- and SAP30L-MO2-treated embryos at 24 hpf compared to RC-treated embryos. ppox mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. The morpholino concentrations used in A and C: 400 μM for MO1, 600 μM for MO2, and 600 μM for RC.

CONCLUSION

SAP30 and SAP30L, members of the well-known Sin3A corepressor complex, have been previously implicated in repression of gene expression. However, the function(s) of these proteins in the setting of a living organism has not been established. Zebrafish was chosen as a model organism because it has only one of these proteins, SAP30L, enabling us to study its role without interference of potentially overlapping functions. Here we have demonstrated that SAP30L is ubiquitously expressed in the zebrafish and participates in the regulation of transcription of multiple genes. Morpholino-mediated knockdown of SAP30L protein results in disruption of cardiac development and function, and a reduction in hemoglobinization of red blood cells. Our results give important in vivo evidence of the role and function of SAP30L during the zebrafish embryogenesis.

ACKNOWLEDGMENTS

We are grateful to Mrs. Leena Mäkinen, Ms. Matilda Martikainen, and Ms. Sanna-Kaisa Harjula (IBT, University of Tampere, Finland) for technical assistance and advice. Dr. Annemarie H. Meijer (Leiden University, the Netherlands) is thanked for technical advice concerning WISH. The microarray analysis was carried out in the Finnish Microarray and Sequencing Centre at the Turku Centre for Biotechnology.

The zebrafish work was carried out in the Tampere zebrafish core facility supported by Biocenter Finland, Tampere Tuberculosis

Foundation and Emil Aaltonen Foundation. The funding sources had no involvement in the study.

REFERENCES

Alexandrovich A, Arno M, Patient RK, Shah AM, Pizzey JA, Brewer AC. 2006. Wnt2 is a direct downstream target of GATA6 during early cardiogenesis. Mech Dev 123:297–311.

Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, Hopkins N. 2004. Identification of 315 genes essential for early zebrafish development. Proc Natl Acad Sci USA 101:12792–12797.

Brownlie A, Donovan A, Pratt SJ, Paw BH, Oates AC, Brugnara C, Witkowska HE, Sassa S, Zon LI. 1998. Positional cloning of the zebrafish sauternes gene: A model for congenital sideroblastic anaemia. Nat Genet 20:244–250.

Chen JN, Fishman MC. 1996. Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. Development 122:3809–3816.

Dannenberg JH, David G, Zhong S, van der Torre J, Wong WH, Depinho RA. 2005. mSin3A corepressor regulates diverse transcriptional networks governing normal and neoplastic growth and survival. Genes Dev 19:1581–1595.

David G, Grandinetti KB, Finnerty PM, Simpson N, Chu GC, Depinho RA. 2008. Specific requirement of the chromatin modifier mSin3B in cell cycle exit and cellular differentiation. Proc Natl Acad Sci USA 105:4168–4172.

Davidson AJ, Zon LI. 2004. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. Oncogene 23:7233–7246.

Detrich HW III, Kieran MW, Chan FY, Barone LM, Yee K, Rundstadler JA, Pratt S, Ransom D, Zon LI. 1995. Intraembryonic hematopoietic cell migration during vertebrate development. Proc Natl Acad Sci USA 92:10713–10717.

Eisen JS, Smith JC. 2008. Controlling morpholino experiments: Don't stop making antisense. Development 135:1735–1743.

Farooq M, Sulochana KN, Pan X, To J, Sheng D, Gong Z, Ge R. 2008. Histone deacetylase 3 (hdac3) is specifically required for liver development in zebrafish. Dev Biol 317:336–353.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch R, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J. 2004. Bioconductor: Open software development for computational biology and bioinformatics.

Goss AM, Tian Y, Tsukiyama T, Cohen ED, Zhou D, Lu MM, Yamaguchi TP, Morrisey EE. 2009. Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. Dev Cell 17:290–298.

Goss AM, Tian Y, Cheng L, Yang J, Zhou D, Cohen ED, Morrisey EE. 2011. Wnt2 signaling is necessary and sufficient to activate the airway smooth muscle program in the lung by regulating myocardin/Mrtf-B and Fgf10 expression. Dev Biol 356:541–552.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, Kelsh RN, Furutani-Seiki M, Vogelsang E, Beuchle D, Schach U, Fabian C, Nusslein-Volhard C. 1996. The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. Development 123:1–36.

Korkeamaki H, Viiri K, Kukkonen MK, Maki M, Lohi O. 2008. Alternative mRNA splicing of SAP30L regulates its transcriptional repression activity. FEBS Lett 582:379–384.

Laherty CD, Billin AN, Lavinsky RM, Yochum GS, Bush AC, Sun JM, Mullen JM, Davie JR, Rose DW, Glass CK, Rosenfeld MG, Ayer DE, Eisenman RN. 1998. SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. Mol Cell 2:33–42.

Liao EC, Paw BH, Peters LL, Zapata A, Pratt SJ, Do CP, Lieschke G, Zon LI. 2000. Hereditary spherocytosis in zebrafish riesling illustrates evolution of erythroid beta-spectrin structure, and function in red cell morphogenesis and membrane stability. Development 127:5123–5132.

Lin SL, Li B, Rao S, Yeo EJ, Hudson TE, Nowlin BT, Pei H, Chen L, Zheng JJ, Carroll TJ, Pollard JW, McMahon AP, Lang RA, Duffield JS. 2010. Macrophage Wnt7b is critical for kidney repair and regeneration. Proc Natl Acad Sci USA 107:4194–4199.

Lindfors K, Viiri KM, Niittynen M, Heinonen TY, Maki M, Kainulainen H. 2003. TGF-beta induces the expression of SAP30L, a novel nuclear protein. BMC Genomics 4:53.

Lobov IB, Rao S, Carroll TJ, Vallance JE, Ito M, Ondr JK, Kurup S, Glass DA, Patel MS, Shu W, Morrisey EE, McMahon AP, Karsenty G, Lang RA. 2005. WNT7b mediates macrophage-induced programmed cell death in patterning of the vasculature. Nature 437:417–421.

Mathavan S, Lee SG, Mak A, Miller LD, Murthy KR, Govindarajan KR, Tong Y, Wu YL, Lam SH, Yang H, Ruan Y, Korzh V, Gong Z, Liu ET, Lufkin Tl. 2005. Transcriptome analysis of zebrafish embryogenesis using microarrays. PLoS Genet 1:260–276.

Montgomery RL, Davis CA, Potthoff MJ, Haberland M, Fielitz J, Qi X, Hill JA, Richardson JA, Olson EN. 2007. Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes Dev 21:1790–1802.

Nakamura Y, Weidinger G, Liang JO, Aquilina-Beck A, Tamai K, Moon RT, Warman ML. 2007. The CCN family member Wisp3, mutant in progressive pseudorheumatoid dysplasia, modulates BMP and wnt signaling. J Clin Invest 117:3075–3086.

Pile LA, Spellman PT, Katzenberger RJ, Wassarman DA. 2003. The SIN3 deacetylase complex represses genes encoding mitochondrial proteins: Implications for the regulation of energy metabolism. J Biol Chem 278:37840–37848.

Pillai R, Coverdale LE, Dubey G, Martin CC. 2004. Histone deacetylase 1 (HDAC-1) required for the normal formation of craniofacial cartilage and pectoral fins of the zebrafish. Dev Dyn 231:647–654.

R Development Core Team. 2008. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rajagopal J, Carroll TJ, Guseh JS, Bores SA, Blank LJ, Anderson WJ, Yu J, Zhou Q, McMahon AP, Melton DA. 2008. Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme. Development 135:1625–1634.

Ransom DG, Haffter P, Odenthal J, Brownlie A, Vogelsang E, Kelsh RN, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Mullins MC, Nusslein-Volhard C. 1996. Characterization of zebrafish mutants with defects in embryonic hematopoiesis. Development 123:311–319.

Reamon-Buettner SM, Borlak J. 2010. NKX2-5: An update on this hypermutable homeodomain protein and its role in human congenital heart disease (CHD). Hum Mutat 31:1185–1194.

Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA, Ekker SC. 2007. P53 activation by knockdown technologies. PLoS Genet 3:e78.

Shu W, Jiang YQ, Lu MM, Morrisey EE. 2002. Wnt7b regulates mesenchymal proliferation and vascular development in the lung. Development 129:4831–4842.

Silverstein RA, Ekwall K. 2005. Sin3: A flexible regulator of global gene expression and genome stability. Curr Genet 47:1–17.

Smyth GK. 2005. Limma: Linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Huber W, Irizarry R, editors. Bioinformatics and computational biology solutions using R and bioconductor. NY: Springer. pp 397–420.

Sousa KM, Villaescusa JC, Cajanek L, Ondr JK, Castelo-Branco G, Hofstra W, Bryja V, Palmberg C, Bergman T, Wainwright B, Lang RA, Arenas E. 2010. Wnt2 regulates progenitor proliferation in the developing ventral midbrain. J Biol Chem 285:7246–7253.

Tang R, Dodd A, Lai D, McNabb WC, Love DR. 2007. Validation of zebrafish (*Danio rerio*) reference genes for quantitative real-time RT-PCR normalization. Acta Biochim Biophys Sin (Shanghai) 39:384–390.

Targoff KL, Schell T, Yelon D. 2008. Nkx genes regulate heart tube extension and exert differential effects on ventricular and atrial cell number. Dev Biol 322:314–321.

Thisse C, Thisse B. 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc 3:59–69.

Viiri KM, Korkeamaki H, Kukkonen MK, Nieminen LK, Lindfors K, Peterson P, Maki M, Kainulainen H, Lohi O. 2006. SAP30L interacts with members of the Sin3A corepressor complex and targets Sin3A to the nucleolus. Nucleic Acids Res 34:3288–3298.

Viiri KM, Heinonen TY, Maki M, Lohi O. 2009a. Phylogenetic analysis of the SAP30 family of transcriptional regulators reveals functional divergence in the domain that binds the nuclear matrix. BMC Evol Biol 9:149.

Viiri KM, Janis J, Siggers T, Heinonen TY, Valjakka J, Bulyk ML, Maki M, Lohi O. 2009b. DNA-binding and -bending activities of SAP30L and SAP30 are mediated by a zinc-dependent module and monophosphoinositides. Mol Cell Biol 29:342–356.

Wang H, Gilner JB, Bautch VL, Wang DZ, Wainwright BJ, Kirby SL, Patterson C. 2007. Wnt2 coordinates the commitment of mesoderm to hematopoietic, endothelial, and cardiac lineages in embryoid bodies. J Biol Chem 282:782–791.

Westerfield M. 1995. The zebrafish book: A guide to the laboratory use of zebrafish. Eugene OR: University of Oregon Press.

Yu J, Carroll TJ, Rajagopal J, Kobayashi A, Ren Q, McMahon AP. 2009. A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. Development 136:161–171.

Zhang Y, Sun ZW, Iratni R, Erdjument-Bromage H, Tempst P, Hampsey M, Reinberg D. 1998. SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. Mol Cell 1:1021–1031.