



Disease-Causing Mutations and Rearrangements in Long Non-coding RNA Gene Loci

Marina Aznaourova¹, Nils Schmerer¹, Bernd Schmeck^{1,2,3} and Leon N. Schulte^{1,2*}

¹Institute for Lung Research, Philipps University Marburg, Marburg, Germany, ²Systems Biology Platform, German Center for Lung Research (DZL), Philipps University Marburg, Marburg, Germany, ³Center for Synthetic Microbiology (SYNMIKRO), Philipps University Marburg, Marburg, Germany

OPEN ACCESS

Edited by:

Claes Wahlestedt,
University of Miami, United States

Reviewed by:

Jörg Hackermüller,
Helmholtz Centre for Environmental
Research (UFZ), Germany

Mainá Bitar,
Federal University of Minas Gerais,
Brazil

John Stanley Mattick,
University of New South Wales,
Australia

*Correspondence:

Leon N. Schulte
leon.schulte@uni-marburg.de;
leon.schulte@staff.uni-marburg.de

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 17 January 2020

Accepted: 02 November 2020

Published: 30 November 2020

Citation:

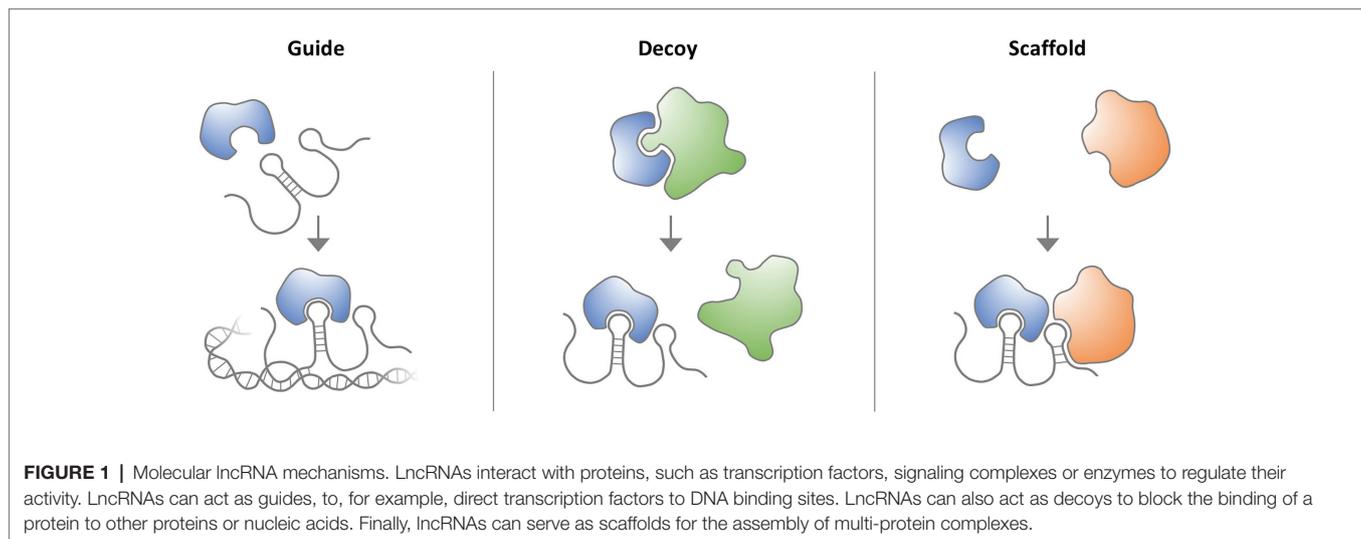
Aznaourova M, Schmerer N,
Schmeck B and Schulte LN (2020)
Disease-Causing Mutations and
Rearrangements in Long Non-coding
RNA Gene Loci.
Front. Genet. 11:527484.
doi: 10.3389/fgene.2020.527484

The classic understanding of molecular disease-mechanisms is largely based on protein-centric models. During the past decade however, genetic studies have identified numerous disease-loci in the human genome that do not encode proteins. Such non-coding DNA variants increasingly gain attention in diagnostics and personalized medicine. Of particular interest are long non-coding RNA (lncRNA) genes, which generate transcripts longer than 200 nucleotides that are not translated into proteins. While most of the estimated ~20,000 lncRNAs currently remain of unknown function, a growing number of genetic studies link lncRNA gene aberrations with the development of human diseases, including diabetes, AIDS, inflammatory bowel disease, or cancer. This suggests that the protein-centric view of human diseases does not capture the full complexity of molecular patho-mechanisms, with important consequences for molecular diagnostics and therapy. This review illustrates well-documented lncRNA gene aberrations causatively linked to human diseases and discusses potential lessons for molecular disease models, diagnostics, and therapy.

Keywords: long non-coding RNA, genome-wide association study, variation (genetic), disease, single nucleotide polymorphism, mutation

INTRODUCTION

An important lesson from the decryption of the human genome sequence and the subsequent systematic mapping of transcribed gene regions is that a surprisingly small proportion of the genome (~2%) encodes proteins. Rather, most of the genome sequence consists of regulatory and structural DNA regions, which harbor thousands of non-coding RNA (ncRNA) loci, including long non-coding RNA (lncRNA) genes (Derrien et al., 2012). lncRNAs are defined as transcripts ≥ 200 nucleotides in length, which do not encode proteins (Derrien et al., 2012) but may regulate the activity of proteins, such as transcription factors or enzymes, by functioning as guides, decoys, or scaffolds (**Figure 1**). lncRNAs can, for instance, act at the chromatin level by forming RNA-DNA hybrids at enhancer and promoter elements and by associating with transcription factors (Rinn and Chang, 2012). Besides guiding and decoying chromatin-regulatory protein complexes, such as the polycomb repressor complexes (PRCs), lncRNAs can also mediate the tethering of distal enhancer elements to gene promoters. Disruption of lncRNA functions consequently contributes to diseases by disturbing key

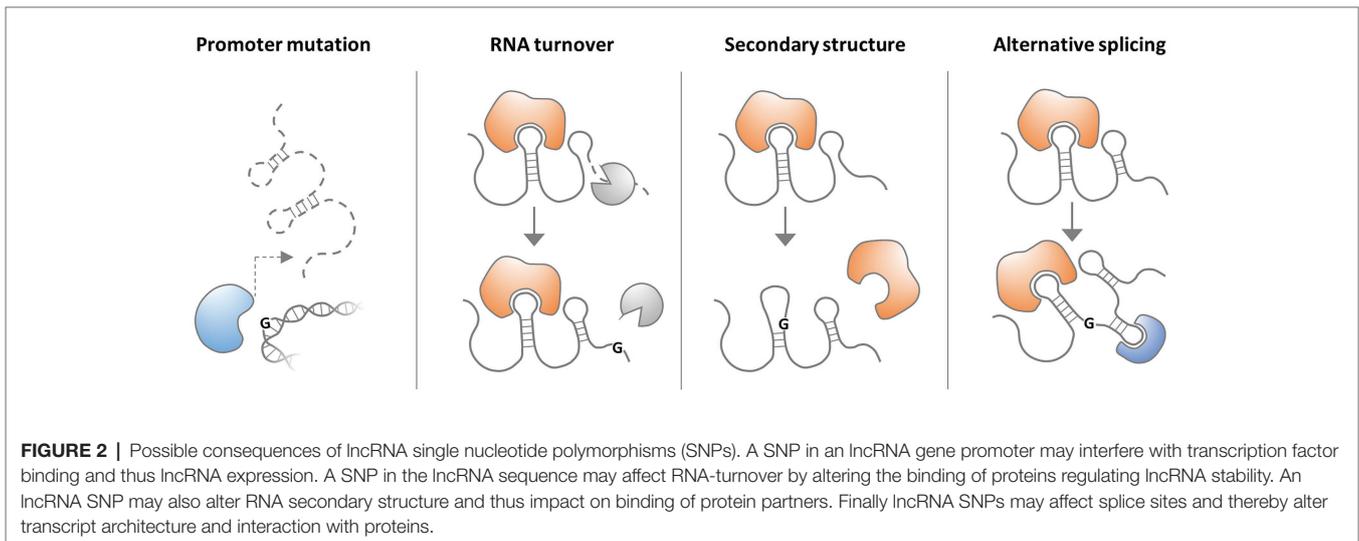


transcriptional circuitries, such as the control of MYC oncogene expression through lncRNA CCAT1-L-associated enhancers (Xiang et al., 2014). Of note, besides regulatory RNA-protein interactions, the act of non-coding RNA transcription itself can adopt important cellular functions (Ali and Grote, 2020). This is exemplified by the prevention of DNA methylation at CTCF binding sites by RNA polymerase II (RNAPolII) mediated ThymoD ncRNA transcription, resulting in a chromatin loop fostering Bcl11b enhancer-promoter interaction (Isoda et al., 2017; section Discussion). Besides the well-established functions in the nucleus, lncRNAs are increasingly recognized to act in cytoplasmic circuitries, where they e.g., regulate translation or mitochondrial function (Carlevaro-Fita and Johnson, 2019). Whereas only few functional lncRNA loci had been discovered until the late 2000s, systematic annotation efforts by the ENCODE and FANTOM consortia have revealed thousands of human lncRNA loci (Derrien et al., 2012; Hon et al., 2017) and numerous recent studies suggest that lncRNAs are critically involved in many human diseases, including Type II diabetes, cancer, or AIDS (Huarte, 2015; Lazar et al., 2016; He et al., 2017). This suggests that the prevalent protein-centric understanding of human diseases does not capture the full complexity of molecular patho-mechanisms, with important consequences for patient genotyping and personalized therapy approaches.

While lncRNA research is still at its infancy, many genome-wide association studies (GWASs) have implicated lncRNA gene polymorphisms in human diseases. The vast majority of GWAS-identified single nucleotide polymorphisms (SNPs) map to intergenic and intronic sequences (Hindorff et al., 2009), supporting the notion that the non-coding regions of the genome are critically involved in disease predisposition. Non-coding SNPs were found to be particularly enriched in promoters, including DNaseI hypersensitive sites (HSSs; Hindorff et al., 2009; Maurano et al., 2012), which represent hallmarks of active transcription and are required for lncRNA gene expression (Janga et al., 2018). Consequently, recent evidence supports

the concept that SNPs may affect the expression of disease-associated lncRNAs through alteration of regulatory DNA regions, such as transcription factor binding sites (Figure 2; Kulkarni et al., 2019). Other consequences of nucleotide rearrangements in disease-relevant lncRNA loci, such as altered secondary structures or alternative splicing, are conceivable (Figure 2). Besides SNPs, long-range alterations, such as amplifications, deletions, and translocations were reported to cause diseases through lncRNA copy number alterations (CNAs) and by affecting lncRNA integrity. Furthermore, several lncRNAs with disease-associated gene variants were proposed as clinical markers, facilitating e.g., stratification and survival prognosis of cancer patients (Zhang et al., 2017; Arriaga-Canon et al., 2018).

Despite the wealth of recently published lncRNA literature, including GWAS and prognostic marker studies, the *in vivo* relevance of lncRNAs in human diseases remains debated, due to contradictory results from cell culture, small animal, and clinical studies. MALAT1, for instance, is abundantly expressed in a variety of human cell types and was among the first lncRNAs to be mechanistically characterized. It is an important structural component of nuclear speckles and has been linked to cancer cell proliferation and metastasis (Zhang et al., 2017c). These results are contrasted by the lack of obvious phenotypes in MALAT1 deficient mice (Eissmann et al., 2012). Whereas cancer-associated polymorphisms have been recurrently identified within the MALAT1 gene locus (Fujimoto et al., 2016; Rheinbay et al., 2020), recent studies cast doubt on the ability of MALAT1 polymorphism data, alone, to establish a causative relationship between the lncRNA and human disease (Carlevaro-Fita et al., 2020; Rheinbay et al., 2020). Additional inclusion of somatic mutation data from tumor genomics or evidence from transposon-mutagenesis screens in mice may pinpoint cancer-driver lncRNAs, including e.g., NEAT1 (Carlevaro-Fita et al., 2020). Presently, however, only few such well-curated resources and meta-analysis pointing out causal associations of lncRNA gene aberrations with diseases are available. Specialized



databases, such as LincSNP 3.0, LncRNADisease 2.0, or NONCODE (Table 1), currently serve as the major knowledge hubs, listing hundreds of reported lncRNA disease associations and polymorphisms, reiteration of which would be beyond the scope of the present review. Rather, we aim at providing an overview of documented disease-causative human lncRNA gene aberrations, which have been genetically characterized and independently confirmed or mechanistically dissected (Table 2). We discuss the utility of such knowledge in molecular diagnostics, patient stratification, and prospective personalized therapy approaches.

DISEASE-CAUSING GENOMIC REARRANGEMENTS AFFECTING lncRNA LOCI

Genomic rearrangements, such as deletions, amplifications, or translocations may alter gene expression, e.g., through the deletion of regulatory DNA elements, or the fusion of silent genes to active enhancer or promoter regions. Such aberrations are key events in the development of prevalent diseases such as cancer, but also in rare Mendelian disorders. Of note, genomic rearrangements causing disease (driver events) need to be discriminated from passenger rearrangements, which constitute a consequence rather than the cause of a disease. Ambitious whole genome sequencing (WGS) projects, such as the 100,000 genome consortiums, and the advent of WGS in the clinics promise to systematically uncover such disease-promoting genomic alterations and to transform molecular diagnosis and patient stratification (Siva, 2015; Nik-Zainal et al., 2020). So far, however, genome-based diagnostics predominantly focuses on protein-coding genes, and the consequences of lncRNA gene disruptions are only beginning to be appreciated. Several studies have implicated deletions, amplifications, and translocations within lncRNA loci in diseases, such as cancer, schizophrenia, or muscular dystrophy.

Section Disease-Causing Genomic Rearrangements Affecting lncRNA Loci summarizes lncRNA gene rearrangements causatively linked to human diseases. Disease-associated amplifications, deletions, and translocations affecting lncRNA genes are complemented by a plethora of SNPs, many of which have been causally linked to disease manifestation and are covered in section SNPs in lncRNA Loci Causally Implicated in Disease.

Amplifications and Deletions

Deletions and amplifications may occur through chromosomal crossover and repair events or replication defects in somatic cells and in the germ-line. While severe germ-line gene defects may stall fetal development (Kacprzak et al., 2016), deletion, or amplification of regulatory DNA elements in proximity to critical genes such as cell-division checkpoint genes may alter their transcript output and predispose to diseases such as cancer. Besides inheritable predispositions, somatically acquired defects in cell division and differentiation genes, e.g., through UV-light exposure or inhalation of mutagens like cigarette smoke, are major drivers of cancer. The molecular diagnostic identification of causative somatic CNAs (SCNAs) due to amplifications or deletions may help to determine appropriate treatment regimens (Nik-Zainal et al., 2020). Recently, besides classic protein-coding genes, such as *MYC*, *RB*, or *p53*, amplifications and deletions affecting lncRNA loci were shown to contribute to oncogenesis (Figure 3). For instance, germ-line deletions of lncRNA ANRIL, located in the *CDKN2A/B* or *INK4-ARF* locus on chromosome arm 9p21, have been associated with an increased risk of melanoma and neural system tumor development in several families in the United States and Europe (Bahuaui et al., 1998; Pasmant et al., 2007). Multiple GWAS studies have confirmed a role of ANRIL in cancer and other diseases (section SNPs Affecting lncRNA Genes Implicated in Oncogenesis). Meanwhile, further lncRNA disruptions due to amplifications and deletions causatively linked to cancer and other diseases have been identified.

TABLE 1 | Databases listing long non-coding RNA (lncRNA) gene variants and disease-associations.

Database	Description and URL	Reference
LincSNP	Disease and phenotype associated gene variants. http://bio-bigdata.hrbmu.edu.cn/lincsnp/index.jsp	Ning et al., 2017
LncVar	SNPs, eQTLs, and copy number variations. http://159.226.118.31/LncVar/	Chen et al., 2017
LncRNADisease	Experimentally supported and predicted disease associations. http://www.rnanut.net/lncrnadisease/index.php/home	Bao et al., 2019
EVLncRNAs	Experimentally validated disease associations. http://biophy.dzu.edu.cn/EVLncRNAs/	Zhou et al., 2018
CLC	List of lncRNAs causally implicated in cancer. https://www.gold-lab.org/clc	Carlevaro-Fita et al., 2020
Lnc2Cancer	Experimentally supported associations with cancer. http://www.bio-bigdata.com/lnc2cancer/	Gao et al., 2019
LncBook	Knowledge database (includes disease-associations and SNPs). https://bigd.big.ac.cn/lncbook/index	Ma et al., 2019
NONCODE	Knowledge database (includes disease-associations and SNPs). http://www.noncode.org/	Fang et al., 2018
LncRNAWiki	Platform for community curation of lncRNA information (including disease associations). http://lncrna.big.ac.cn/index.php/Main_Page	Ma et al., 2015

TABLE 2 | lncRNAs covered in the present review, genetically linked to human diseases.

lncRNA	Mechanism	Disease	Reference
ALAL-1	<i>trans</i> Interacts with SART3 and regulates USP4 deubiquitinase subcellular shuttling	Adenocarcinoma	Athie et al., 2020
ANRIL	<i>cis</i> and <i>trans</i> Regulation of CDKN2A/B <i>via</i> binding to CBX7, interaction with YY1 transcription factor	Melanoma, breast cancer, neuronal system tumor, type 2 diabetes, glaucoma, cardiovascular diseases	Helgadottir et al., 2007; McPherson et al., 2007; Pasmant et al., 2007; Scott et al., 2007; Zeggini et al., 2007; Broadbent et al., 2008; Stacey et al., 2009; Bei et al., 2010
ATXN80S	?	Spinocerebellar ataxia, Parkinson disease	Koob et al., 1999; Samukawa et al., 2019
CCAT2	<i>trans</i> Induction of expression of WNT signaling genes, regulation of MYC expression, and alternative splicing of GLS mRNA	Cancer	Pomerantz et al., 2009; Ling et al., 2013; Redis et al., 2016; Xin et al., 2017
CCR5AS	<i>cis</i> Enhancement of CCR5 mRNA stability <i>via</i> suppressive interaction with RALY	HIV infection	McLaren et al., 2015; Kulkarni et al., 2019
CUPID1/2	<i>trans</i> Locate to nuclear foci and promote DNA repair by homologous recombination	Breast cancer	Betts et al., 2017
DA125942	<i>cis</i> PTH1LH expression	BDE	Maass et al., 2012
DBE-T	<i>cis</i> Recruitment of Ash1L to FSHD gene locus	FSHD	Cabianca et al., 2012
DIRC3	<i>trans</i> promotes expression of the tumor suppressor IGFBP5	Renal carcinoma, melanoma	Bodmer et al., 2003; Coe et al., 2019
DISC2	?	Schizophrenia	Millar et al., 2000; Blackwood et al., 2001; Chubb et al., 2008
FAL1	<i>trans</i> Stabilization of BMI1	Ovarian cancer, prostate cancer	Hu et al., 2014
Gas5	<i>cis</i> and <i>trans</i> Fusion with BCL6, decoy of the glucocorticoid receptor	B-cell lymphoma	Nakamura et al., 2008
HOTAIR	<i>trans</i> Regulation of HOX gene expression	Cancer	Rinn et al., 2007; Tang and Hann, 2018
HULC	<i>trans</i> Interaction with YB1, coREST, and hnRNPK	Cancer, hepatocellular carcinoma, colorectal cancer, esophageal cancer	Liu et al., 2012; Kang et al., 2015; Shaker et al., 2017; Yu et al., 2017; Chen et al., 2019b
H19	<i>trans</i> Interaction with chromatin factors let-7 microRNA sponge	Atherosclerosis, coronary artery disease, predisposition to elevated blood-pressure, cancer	Han et al., 1996; Verhaegh et al., 2008; Kallen et al., 2013; Monnier et al., 2013; Tragante et al., 2014; Gao et al., 2015; Sung et al., 2018; Li et al., 2018a
IPW	<i>trans</i> Reduction of gene expression in DLK1-DIO3 gene locus	PWS	Wevrick et al., 1994

(Continued)

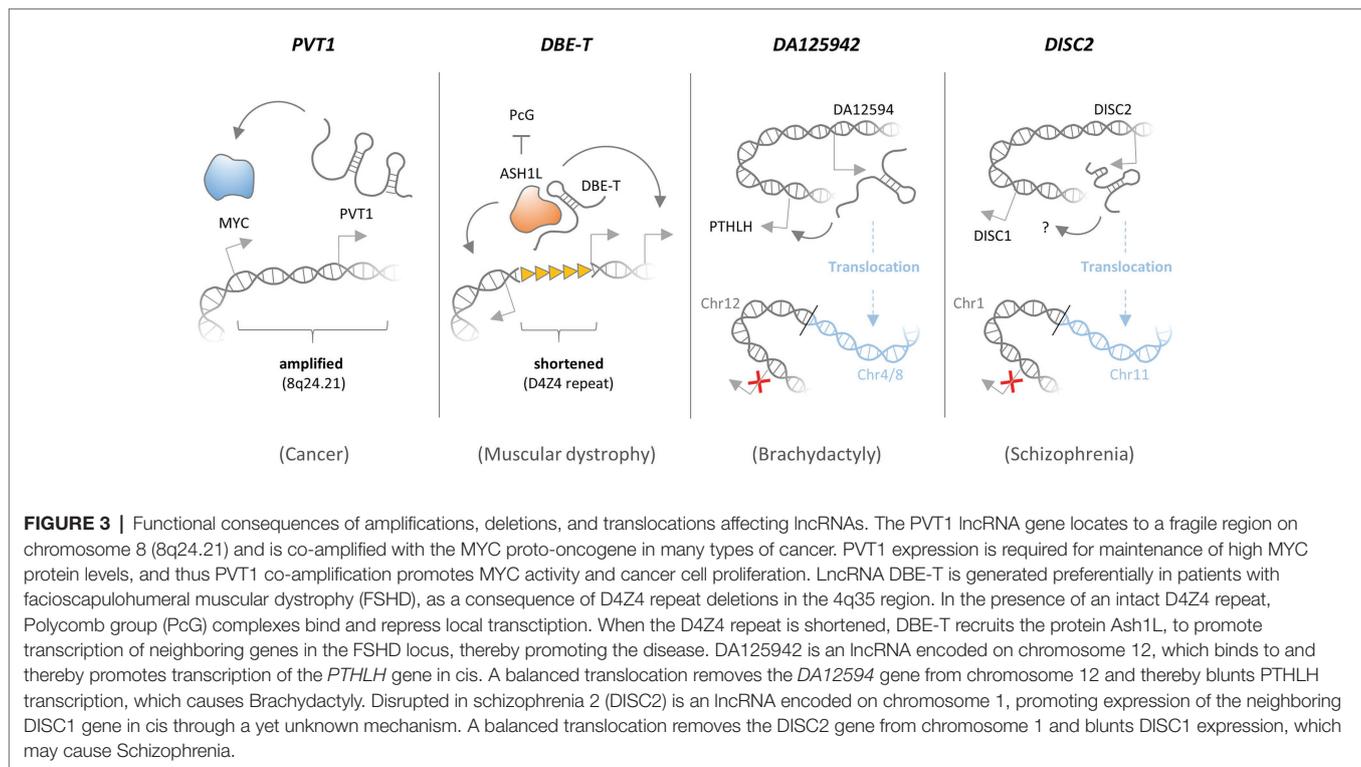
TABLE 2 | Continued

LncRNA	Mechanism	Disease	Reference
IFNG-AS	<i>cis</i> Interaction with WDR5, promotion of IFN γ expression	IBD	Vigneau et al., 2003; Mirza et al., 2015; Padua et al., 2016
Lnc-NR2F1	<i>trans</i> Binds to chromatin and regulates neuronal genes	Developmental disorders	Ang et al., 2019
Lnc13	<i>cis</i> and <i>trans</i> Repression of inflammatory gene transcription via interaction with hnRNP D	Celiac disease	Castellanos-Rubio et al., 2016
LINC00237	?	MOMO syndrome	Vu et al., 2012
LINC00305	?	Atherosclerosis, rheumatoid arthritis	Zhang et al., 2017b; Wahba et al., 2020
Linc-HELLP	<i>trans</i> Alteration of protein interactions involved in splicing and ribosomal function	HELLP syndrome	van Dijk et al., 2012, 2015
LOC285194	?	Osteosarcoma	Li et al., 2017b; Pasic et al., 2010
MIAT	<i>trans</i> Interacts with the nuclear matrix and associates with splice factor SF1	Cardiovascular diseases	Ishii et al., 2006; Sone et al., 2007; Tsuji et al., 2011; Ishizuka et al., 2014
MIR2052HG	<i>trans</i> Promotes ER α expression and stability	Breast cancer	Ingle et al., 2016
NEAT1	?	Cancer, breast cancer, neural degeneration	Fujimoto et al., 2016; Atianand et al., 2017; Rheinbay et al., 2017; An et al., 2018; Li et al., 2018b
OVAL	?	Ovarian adenocarcinoma and endometrial carcinoma	Akrami et al., 2013
PCAN-R1 and PCAN-R2	?	Prostate cancer	Du et al., 2013
PCAT1	<i>trans</i> Regulation of androgen-stimulated genes via interaction with AR/LSD1 complex	Prostate cancer	Prensner et al., 2011
PRAL	<i>trans</i> Regulates p53 stability	hepatocellular carcinoma	Zhou et al., 2016a
PTCSC3	?	papillary thyroid carcinoma	Jendrzewski et al., 2012
PVT1	<i>cis</i> and <i>trans</i> Stabilization of MYC, KLF5 and STAT3	Cancer, small cell lung carcinoma	Pleasant et al., 2010; Tseng et al., 2014; Tang et al., 2018; Zhao et al., 2018; Jin et al., 2019
RMRP	<i>trans</i> rRNA maturation?	CHH, cancer	Martin and Li, 2007; Zack et al., 2013; Rheinbay et al., 2020
RMST	<i>trans</i> Associates with hnRNPA2/B1 and SOX2 to promote SOX2 target gene expression	Kallmann syndrome	Ng et al., 2013; Stamou et al., 2020
RUNXOR	<i>trans</i> engaged in chromatin loops, relevant to RUNX1 translocations	Acute myeloid leukemia	Wang et al., 2014
SAMMSON	<i>trans</i> Regulation of mitochondrial metabolism via interaction with p32	Melanoma	Leucci et al., 2016
SNHG17	?	Cancer	Ma et al., 2017; Xu et al., 2019

NEAT1

The essential nuclear speckle component NEAT1 belongs to the most well-studied lncRNAs. Besides its role as an architectural RNA of nuclear bodies, NEAT1 is also involved in inflammatory, neurodegenerative, and other disease-associated cellular processes (Atianand et al., 2017; An et al., 2018). Furthermore, NEAT1 is over-expressed in many solid tumors (Li et al., 2018b) and knockdown of NEAT1 decreased cell invasion (Fujimoto et al., 2016). Multiple recent reports suggest a causal link between NEAT1

gene aberrations and cancer. NEAT1 was identified as one out of three genes with recurrent mutations in breast cancer. Interestingly, three out of four identified promoter mutations interfered with NEAT1 expression, in line with its frequent focal deletion in breast cancers (Rheinbay et al., 2017). Thus, both NEAT1 gain- and loss-of-function might contribute to oncogenesis. In a pan-cancer study, an enrichment of indel mutations in the NEAT1 locus was observed. It was doubted, however, that these constitute driver mutations (Rheinbay et al., 2020). In another



recent study, including information from transposon-mutagenesis screens, however, NEAT1 was identified as one out of eight lncRNAs with cancer driver mutations (Carlevaro-Fita et al., 2020). Different from other well-studied lncRNAs, such as PVT1 or ANRIL (see below), the precise molecular mechanisms by which NEAT1 impacts on oncogenesis remain to be determined. Taken together, the well-studied RNA component of nuclear speckles, NEAT1, is recurrently mutated and subjected to SCNAs, causally implicated in cancer.

PVT1

8q24.21 is a fragile region on chromosome 8, which is amplified in many types of cancer and harbors the MYC proto-oncogene (Jin et al., 2019). Increased MYC expression has been implicated in the development of many types of tumors. Consequently, protective cellular mechanisms exist, which induce growth arrest or cell death in response to elevated MYC expression (Gabay et al., 2014). Thus, MYC activation alone is typically not sufficient to reinforce cell transformation programs and initiate tumorigenesis. Besides the MYC gene, the 8q24.21 region harbors the PVT1 lncRNA locus, expression of which is consequently elevated in various types of cancer tissues and cell lines as a result of amplifications and translocations (Boloix et al., 2019; Jin et al., 2019). A report by Tseng et al. (2014) suggests that the gain of PVT1 lncRNA expression is required for the maintenance of high MYC protein levels and cancer cell proliferation. Besides MYC, PVT1 was also shown to stabilize other cell proliferation and survival promoting proteins, such as KLF5 and STAT3 (Tang et al., 2018; Zhao et al., 2018). Consequently, PVT1 was identified as a prognostic marker for

cancer patient's overall survival, and elevated PVT1 expression was found to associate with advanced tumor severity stage according to the tumor-node-metastasis (TNM) classification scheme (Lu et al., 2017; Martini et al., 2017; He et al., 2018; Boloix et al., 2019). Interestingly, two alternative PVT1 transcriptional start sites were shown to serve as enhancer elements, directly regulating MYC gene expression independent of PVT1 lncRNA expression (Fulco et al., 2016). Thus, the PVT1 lncRNA locus, which is often co-amplified with the MYC gene, may contribute to tumorigenesis through both lncRNA-mediated MYC protein stabilization and an lncRNA independent enhancer function (Figure 3).

FAL1

An array- and shRNA-based screen across tumor tissues from various types of cancer pinpointed FAL1 as an oncogenic lncRNA expressed from a focal amplicon on chromosome 1q21.2 (Hu et al., 2014). FAL1 copy number gain significantly associated with decreased ovarian cancer and prostate cancer patient's overall survival (Hu et al., 2014; Zhao et al., 2017). In melanoma patients, FAL1 expression was shown to have prognostic value with respect to lymph node metastasis and TNM stage, as well as overall survival (Ni et al., 2017). Altered expression and prognostic potential of FAL1 have also been reported in the context of diverse other types of cancer (Lv et al., 2019). Mechanistically, FAL1 associates with and stabilizes the PRC 1 component BMI1 in a ubiquitin-proteasome dependent manner to suppress expression of the cyclin-dependent kinase inhibitor p21 (Hu et al., 2014). Thus, similar to NEAT1 and PVT1, CNAs affecting lncRNA FAL1 are causally involved in oncogenesis and have prognostic potential.

RMRP

Long non-coding RNA RMRP is a component of the mitochondrial RNA processing endoribonuclease MRP, which is involved in rRNA maturation (Martin and Li, 2007). Germ-line point-mutations, insertions, and duplications reducing RMRP expression cause the recessively inherited disorder “cartilage-hair hypoplasia” (CHH), which is primarily associated with abnormalities in bone and cartilage development (Ridanpaa et al., 2001; Martin and Li, 2007; Nakashima et al., 2007). Besides CHH, several cancer-genome studies have revealed RMRP SCNAs due to focal amplifications (Zack et al., 2013; Rheinbay et al., 2020), and the RMRP promoter was found to contain a breast-cancer associated mutation hotspot, impacting RMRP transcription (Rheinbay et al., 2017). In tissue samples from other types of cancer, RMRP mutations were identified with only low prevalence (0–1.3%), suggesting that RMRP SCNAs do not constitute generalized cancer drivers or biomarkers, beyond the well-documented implications in breast cancer (Son et al., 2019). Thus, loss and gain of lncRNA RMRP expression due to point-mutations and amplifications is causally linked to CHH and breast cancer, respectively.

SAMMSON

Similar to RMRP, SAMMSON is a nuclear encoded lncRNA localized to mitochondria. It promotes mitochondrial metabolism by fostering mitochondrial localization of rRNA processing factor p32 (Leucci et al., 2016). In melanoma patients, the SAMMSON locus is recurrently co-gained with the MITF oncogene through 3p13-3p14 focal amplification. Additionally, SAMMSON expression in melanoma cells is promoted by the SOX10 transcription factor. SAMMSON promotes melanoma cell viability and desensitizes the cells to MAPK-targeting cancer therapeutics (Leucci et al., 2016). Thus, both RMRP and SAMMSON are mitochondrial metabolic regulators with recurrent SCNAs due to focal amplifications, causally involved in cancer. Of note, MITF, with which SAMMSON is co-amplified at 3p13-3p14, is a transcriptional regulator of lncRNA DIRC3, which is disrupted by balanced translocations in renal carcinoma and suppressed by SOX10 in melanoma cells (see Translocations).

DBE-T

Facioscapulohumeral muscular dystrophy (FSHD) is an inherited disease, which is characterized by progressive skeletal muscle weakness. The disease has been linked to a shortened D4Z4 repeat in the subtelomeric chromosome region 4q35, which controls expression of the FSHD gene locus in cis (Sacconi et al., 2015). The FSHD locus may be silenced by PRC proteins, which bind to the D4Z4 repeat and establish repressive chromatin marks, or activated by Trithorax proteins during myoblast differentiation (Bodega et al., 2009; Cabianca et al., 2012). In FSHD patients, shortening of the D4Z4 repeat region was found to activate the FSHD locus through an epigenetic mechanism involving the cis-acting lncRNA DBE-T. DBE-T is generated preferentially in FSHD patients, as a consequence of D4Z4 repeat deletions and recruits the Trithorax protein Ash1L, to promote H3K36 dimethylation and transcription of

the FSHD locus (Cabianca et al., 2012; **Figure 3**). Thus, increased expression of nuclear lncRNA DBE-T due to D4Z4 repeat shortening induces chromatin-changes at the site of lncRNA transcription, triggering expression of neighboring disease-associated transcripts.

IPW

Imprinting controls the exclusive expression of genes from the paternal or maternal chromosome. Microdeletions in the imprinted 15q11-q13 region on the paternal chromosome cause Prader-Willi Syndrome (PWS), a genetic disease associated with life-threatening obesity, but also learning and behavioral problems (Butler et al., 2019). Already in 1994, the PWS-associated chromosomal region was shown to give rise to the IPW lncRNA, which is lost in PWS patients (Wevrick et al., 1994). Encoded on chromosome 15, IPW acts in *trans* to reduce expression of genes from the DLK1-DIO3 locus on maternal chromosome 14. Thus, loss of trans-regulation of DLK1-DIO3 expression by IPW might contribute to PWS (Stelzer et al., 2014). IPW spans the SNORD116 snoRNA cluster, and deletions affecting these snoRNAs were implicated in PWS in humans (Sahoo et al., 2008). In mice, the lncRNA host gene giving rise to SNORD116 snoRNAs was shown to interact with transcriptional regulator RBBP5 in the nucleus. Deletion of the murine SNORD116 host-gene disrupted expression of circadian clock genes and caused altered diurnal energy expenditure in the brain, which might explain some of the neurological problems in PWS patients (Powell et al., 2013). Taken together, clinical and genetic evidence collected over the past 25 years pinpoints disruption of the 15q11-q13 region as a common cause of PWS, rendering it one of the first lncRNA loci linked to a human genetic disease. The precise mechanistic contributions of 15q11-q13 encoded ncRNAs to PWS still remain to be determined.

Besides the well-documented amplifications and deletions illustrated above, several other lncRNA CNAs have been reported and linked to human diseases. For instance, an expanding tri-nucleotide repeat within the ATXN8OS lncRNA locus silences ATXN8OS expression and has been associated with spinocerebellar ataxia and Parkinson disease (Koob et al., 1999; Samukawa et al., 2019). ALAL-1, an immune-inducible lncRNA, is amplified in adenocarcinomas and promotes tumorigenesis (Athie et al., 2020). lncRNA LOC285194 is recurrently deleted in osteosarcoma and associated with poor survival (Pasic et al., 2010; Li et al., 2017b). Focal amplification of the OVAL lncRNA locus on chromosome 1 was observed in ovarian adenocarcinoma and endometrial carcinoma (Akrami et al., 2013). In a screen for cancer driver lncRNAs, PCAN-R1 and PCAN-R2 were identified, which locate to SCNA regions and display increased expression in prostate cancer tissue. Knockdown of both lncRNAs slowed down prostate cancer cell growth *in vitro* (Du et al., 2013). Similarly, SNHG17, which is overexpressed in different cancer tissues, likely through amplifications, promotes cancer cell proliferation and constitutes a marker of poor prognosis (Ma et al., 2017; Xu et al., 2019). PRAL is a recurrently deleted

lncRNA in hepatocellular carcinoma, which regulates p53 stability (Zhou et al., 2016a). Additional, lncRNAs are likely subjected to CNAs and constitute potential cancer drivers, as suggested by systematic screens (Zhou et al., 2017; Volders et al., 2018; Luo et al., 2019; Athie et al., 2020). While dozens of additional cancer-associated lncRNAs exist (Du et al., 2013; Carlevaro-Fita et al., 2020), studies systematically linking lncRNA CNAs due to amplifications and deletions to clinical outcomes, and disease mechanisms are still scarce. The examples illustrated above, however, suggest that further research into the clinical relevance of lncRNA CNAs may reveal additional driver mutations, valuable prognostic markers, and potential novel therapeutic angles.

Translocations

Besides amplifications and deletions, translocations may affect the functionality and the transcriptional output of disease-relevant lncRNA genes. The term “translocation” describes an unusual chromosomal rearrangement, e.g., caused by a chromosomal break, which is followed by fusion of the resulting fragment to a different chromosome. Balanced translocations, in which no genetic material is lost, may occur without causing disease symptoms. Sometimes, however, the function of disease-relevant genes is altered. Translocations may occur non-randomly at chromosomal break-points, and many of the resulting malignancies have been characterized (Aplan, 2006). lncRNA genes (e.g., lnc-RP11-211G3.3.1-1 or FAM230C) may demarcate such chromosomal break-points (Lu et al., 2015; Delihias, 2018) or be fused to regulatory elements as a result of translocations, thereby causing disease. For instance, fusion of the 5' region of the PVT1 lncRNA gene to the CHD7 chromatin remodeler gene likely results in CHD7 gain-of-function in small cell lung cancer (Pleasant et al., 2010). Dozens of further cancer-associated PVT-1 translocation events have been documented (Jin et al., 2019), underpinning the reported roles of this lncRNA in oncogenesis (section Amplifications and Deletions). Additional, translocation events disturbing disease-associated lncRNA loci have been described (Figure 3). Well-documented examples are depicted below.

Growth Arrest Specific Transcript 5

Growth arrest specific transcript 5 (GAS5) was originally identified as a serum-starvation induced transcript (Schneider et al., 1988), which functions as a decoy of the glucocorticoid receptor (Kino et al., 2010). Fusion of the GAS5 and the BCL6 oncogene as the consequence of a $t(1;3)(q25;q27)$ translocation was reported in a patient with B-cell lymphoma (Nakamura et al., 2008). While there is presently no indication for a broader relevance of this fusion event in cancer, GAS5 has in the meantime been confirmed as a transcript with broad implications in oncogenesis and with prognostic potential for various cancers (Ji et al., 2019).

DIRC3

Another lncRNA disrupted by a documented cancer-relevant translocation event is DIRC3. In a family with renal cell cancer,

the $t(2;3)(q35;q21)$ translocation was identified, and several genes were mapped to the chromosomal break-point (Bodmer et al., 2002, 2003). The translocation results in a fusion of the DIRC3 lncRNA gene with the HSPBAP1 gene, with yet to be determined mechanistic consequences. In a screen for lncRNAs targeted by the melanoma transcription factors MITF and SOX10, DIRC3 was identified as a top hit and characterized as a melanoma tumor suppressor (Coe et al., 2019). Mechanistically, DIRC3 seems to promote expression of the tumor suppressor IGFBP5 through a chromatin loop, which brings both genes into close proximity. MITF and SOX10 suppress expression of DIRC3, which potentially contributes to melanoma development (Coe et al., 2019). Of note, MITF is co-gained with lncRNA SAMMSON (see section Amplifications and Deletions), suggesting several lncRNAs to act in a common cancer-associated circuitry.

Disrupted in Schizophrenia 2

Besides cancer, lncRNA-affecting translocations have also been linked to psychiatric and developmental disorders. The “disrupted in schizophrenia” genes *DISC1* and *DISC2*, for instance, map to a locus on chromosome 1, where they are transcribed into opposite directions. Whereas *DISC1* encodes a peptide, *DISC2* is transcribed into a lncRNA molecule and overlaps with *DISC1* (Figure 3). The DISC locus resides within a chromosomal breakpoint region and is disrupted by the balanced translocation $t(1;11)(q42;q14.3)$, which co-segregates with schizophrenia and affective disorders (Millar et al., 2000; Blackwood et al., 2001). *DISC1* functions as a signaling hub, which aggregates proteins involved in pathways relevant to neuronal function and psychiatric illness (Chubb et al., 2008). The molecular function of schizophrenia associated lncRNA *DISC2* remains to be clarified.

DA125942

Brachydactyly Type E (BDE) is a genetic disorder, characterized by shortened metacarpal and metatarsal bones. Mutations, in the *HOXD13* and parathyroid hormone-like hormone (*PTH1LH*) genes have been linked to the disease (Klopocki et al., 2010). *PTH1LH* is an important regulator of chondrogenesis, the process of cartilage formation and condensation (Stricker and Mundlos, 2011). In BDE affected individuals, balanced translocations affecting chromosome 12p were identified, which disrupt a cis-regulatory element required for normal *PTH1LH* expression (Maass et al., 2012). This element was found to be transcribed and give rise to an lncRNA, DA12594. In normal cells, DA12594 lncRNA was found to physically associate with chromatin at the *PTH1LH* gene and to be required for *PTH1LH* transcription. The authors proposed that the identified translocations in BDE patients disrupt the regulatory interaction between the DA12594 lncRNA locus and the *PTH1LH* gene, thereby causing reduced *PTH1LH* expression and disturbed chondrogenesis (Maass et al., 2012; Figure 3).

LINC00237

Genetic obesity syndromes and developmental disorders may be caused by a range of different genomic aberrations, including

the IPW lncRNA gene disruption in PWS (see section Amplifications and Deletions). MOMO syndrome is another rare, genetic disorder, characterized by macrosomia, obesity, macrocephaly, and ocular abnormalities, as well as mental disability (Moretti-Ferreira et al., 1993; Di Donato et al., 2012). In a patient diagnosed with MOMO syndrome, Vu et al. (2012) identified the homozygous balanced translocation t(16;20)(q21;p11.2). The breakpoint at 20p11.23 disrupted the ncRNA gene LINC00237 and thereby, the t(16;20)(q21;p11.2) translocation interfered with expression of this lncRNA (Vu et al., 2012). In summary, the molecular dissection of the t(16;20)(q21;p11.2) translocation revealed yet another lncRNA causatively linked to a human developmental disorder; the precise molecular function of LINC00237 in human cells, however, remains to be determined.

RMST

Another rare developmental disorder is the Kallmann syndrome (KS), which is characterized by a disturbed gonadotropin-releasing hormone (GnRH) balance, and thus delayed sexual maturation, combined with anosmia. Several gene defects resulting in GnRH deficiency could be attributed to KS. Of note, in a KS patient, the balanced translocation t(7;12)(q22;q24) was identified, disrupting the RMST lncRNA gene and causing reduced RMST expression in GnRH-targeted neurons. Furthermore, expression of several genes associated with the GnRH pathway was reduced (Stamou et al., 2020). RMST was previously shown to regulate neuronal differentiation by associating with hnRNP2/B1 and the SOX2 transcription factor, thereby promoting SOX2 target gene expression (Ng et al., 2013). Together, these reports suggest that RMST disruption contributes to KS by affecting proper GnRH-dependent neuronal maturation at the chromatin level.

Additional, lncRNA genes were shown to be affected by translocation events. Lnc-NR2F1, for instance, is another lncRNA implicated in neurogenesis and involved in developmental disorders. This lncRNA was found to be disrupted by a t(5;12) translocation in a family with neurodevelopmental symptoms. Lnc-NR2F1 binds to chromatin and regulates neuronal genes. In the same study, additional lncRNAs were found to be recurrently mutated in patients with intellectual disability and autism spectrum disorders (Ang et al., 2019). Other lncRNAs, demarcating chromosomal translocation break points exist (Delihias, 2018) and remain to be mechanistically dissected or causally implicated in human diseases. Furthermore, altered activity of signaling pathways and transcription factors due to translocations affecting coding genes impacts the expression of disease-associated classical GENCODE lncRNAs and enhancer RNAs, thereby e.g., contributing to leukemia and drug-resistance (Teppo et al., 2016; Han et al., 2019; Ng et al., 2019). Additionally, lncRNAs may fuse to coding genes to promote their activation. An example is the documented fusion of MALAT1 to GLI1, which enhances GLI1 expression, and thus hedgehog signaling in different types of cancer (Spans et al., 2016; Graham et al., 2017; Antonescu et al., 2018). Finally, lncRNAs might actively contribute to translocation events, as exemplified by

lncRNA RUNXOR. This lncRNA seems to be engaged in chromatin loops, potentially involved in RUNX1 translocations, often observed in AML (Wang et al., 2014).

Taken together, the here depicted examples suggest important contributions of lncRNA-affecting translocations to a variety of human diseases and predict that systematic mapping of disease-causative long-range alterations in lncRNA genes will further improve diagnosis and patient stratification. An example is the facilitated discrimination of Prader-Willi and MOMO syndrome by distinct genomic signatures in lncRNA gene loci. While the discrimination in this case does not necessarily require molecular diagnostics, the latter may further underpin the disease-cause and pave the way for a better understanding of obesity-related genetic diseases and therapeutic angles. It remains to be determined, whether the respective lncRNAs may serve as direct therapeutic targets for tailored therapies (see section Discussion).

SNPs IN lncRNA LOCI CAUSALLY IMPLICATED IN DISEASE

The decryption of the human genome sequence and the increased availability of array and next-generation sequencing technologies have propelled the identification of genetic variants, associated with specific disease traits. Besides large genomic aberrations, focal amplifications and deletions, GWASs have already mapped thousands of disease-relevant SNPs (MacArthur et al., 2017). However, the clinical relevance of GWAS may be limited by the fact that individual SNPs typically account for only a small proportion of the variants underlying heritability of complex traits (Tam et al., 2019). In addition, GWAS approaches often focus on disease indicators, which are easy to quantify (e.g., the body-mass-index) and may thus underestimate disease complexity (Tam et al., 2019). On the other hand, GWAS may pinpoint molecular patho-mechanisms, which can guide the development of novel therapeutics and narrow down personalized treatment options, based on patient genotyping data.

Several possible scenarios can be envisioned, by which SNPs affect the functions of lncRNAs (Figures 2, 4). Disease-associated SNPs are often found in regulatory promoter elements (Hindorff et al., 2009; Maurano et al., 2012), where they may alter the binding of transcription factors, and thus impact on lncRNA expression levels. Single nucleotide variants may also be located in mature lncRNA sequences and affect protein binding motifs, e.g., required for post-transcriptional control of RNA stability. SNPs within lncRNA secondary structures, such as hairpin loops, may alter RNA folding and thereby the interaction with other biomolecules. SNPs may also affect splice-sites, thereby promoting the accumulation of alternative splice-variants with altered functionality (Figure 2). While a complete list of published SNPs within lncRNA loci would be beyond the scope and limits of the current review, the following sections cover well-characterized examples and illustrate mechanistic consequences of lncRNA SNPs in the context of human diseases.

SNPs Affecting lncRNA Genes Implicated in Oncogenesis

Cancer remains a leading cause of death world-wide, and research over the past decade has identified a number of common cancer driver mechanisms. Since cancer may originate from a variety of different tissues and cell types, patients may profit from personalized therapies based on their cell-type specific driver mutations (Nik-Zainal et al., 2020). lncRNAs are generally more cell-type specific than mRNAs and recently, potential driver mutations in lncRNA loci have been revealed (Carlevaro-Fita et al., 2020). Thus, lncRNAs might constitute attractive candidates for prospective personalized therapies (Nguyen and Carninci, 2016). The database LincSNP 2.0 currently lists 371,647 disease-associated SNPs in lncRNA loci, of which ~11% overlap with transcription factor binding sites in lncRNA promoters (Ning et al., 2017). Among the overall disease-associations reported by the lncRNA disease database, lncRNAs most frequently associate with cancer and the same holds true for lncRNA gene polymorphisms (Bao et al., 2019). In the following, well-documented cancer-associated SNPs affecting lncRNA loci and their mechanistic consequences in the context of disease predisposition and pathogenesis are described.

ANRIL

Already before the rise of GWAS, deletions of the ANRIL lncRNA gene upstream of the CDKN2A and CDKN2B genes have been associated with increased risk of melanoma and neural system tumor development in several families (see section Amplifications and Deletions). The ANRIL locus contains three coding genes (*CDKN2A*, *CDKN2B*, and *ARF*) that are critical for the function of the retinoblastoma (RB) and p53 tumor suppressor networks (Sherr, 2012). ANRIL overlaps with the *ARF* and *CDKN2B* genes and is co-expressed with both genes under normal and pathologic conditions (Pasmant et al., 2007). Several GWASs have associated SNPs in this region not only with diverse types of cancer (Stacey et al., 2009; Bei et al., 2010; Turnbull et al., 2010), but also other diseases, such as coronary heart disease (Helgadottir et al., 2007; McPherson et al., 2007; Broadbent et al., 2008), or Type 2 diabetes (Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research et al., 2007; Scott et al., 2007; Zeggini et al., 2007). By now, it has become widely accepted that SNPs in the CDKN2A and CDKN2B upstream region act predominantly through lncRNA ANRIL (Pasmant et al., 2011). In line with this view, SNPs in the CDKN2A/CDKN2B locus were reported to impact ANRIL lncRNA expression (Cunnington et al., 2010). Disease associated SNPs were reported to affect a repressive STAT1 binding site near the ANRIL locus (Harismendy et al., 2011). Derepressed ANRIL in turn negatively regulates the neighboring CDKNB gene (Yap et al., 2010; Harismendy et al., 2011). Mechanistically, ANRIL binds to CBX7, a component of the PRC1 and promotes silencing of CDKNB expression in a PRC1- and histone H3K27me-dependent manner (Yap et al., 2010). Disease SNPs or promoter deletions, negatively impacting on ANRIL expression (Cunnington et al., 2010), might relieve

CDKNB repression through the same mechanism. In line with the inflammatory component in coronary artery disease, which ANRIL SNPs predispose to, ANRIL expression is elevated by pro-inflammatory stimuli, such as the cytokines interferon-gamma (IFN γ) or TNF α (Harismendy et al., 2011; Zhou et al., 2016b). Under these conditions, ANRIL was found to interact with the YY1 transcription factor to promote expression of important mediators of the inflammatory response (Zhou et al., 2016b). Thus, lncRNA ANRIL, expression of which is affected by several disease-associated SNPs, regulates gene expression both in *cis* and in *trans* at the chromatin level (Figure 4). Notably, ANRIL was also described to exist in a circularized form, which controls ribosomal RNA maturation and inhibits cell proliferation (Burd et al., 2010; Holdt et al., 2016). This suggests that ANRIL is a multifunctional lncRNA involved in both transcriptional and post-transcriptional regulations of disease-relevant cellular pathways.

CCAT2

Similar to ANRIL, the CCAT2 lncRNA gene, located in chromosomal region 8q24.21, has been implicated in several malignancies, and CCAT2 expression is increased in many types of cancer (Xin et al., 2017). The CCAT2 gene region contains a colorectal cancer associated SNP (rs6983267), and the risk allele was shown to promote CCAT2 overexpression (Ling et al., 2013). Mechanistically, CCAT2 binds to TCFL2 to promote expression of WNT signaling genes. CCAT2 is itself a WNT signaling target, suggesting that the lncRNA establishes a positive feed-back loop, promoting the activity of this pathway and eventually WNT-dependent MYC proto-oncogene expression (Ling et al., 2013). The CCAT2 SNP (rs6983267) region was also reported to interact with the MYC promoter (Pomerantz et al., 2009), suggesting that the CCAT2 gene might directly impact on MYC expression through an enhancer-like function. The G-allele of the CCAT2 rs6983267 SNP has also been reported to preferentially bind the CFIm25 subunit of the cleavage factor I (CFIm) complex and promote alternative splicing of Glutaminase (GLS) mRNA, to promote cancer cell metabolism and proliferation (Redis et al., 2016). Furthermore, CCAT2 transgenic mice, carrying the disease-SNP alleles develop myeloid malignancies, which can be traced back to impaired PRC2 function (Shah et al., 2018). Thus, CCAT2, similar to ANRIL appears to be a multifunctional lncRNA, involved in cancerogenesis through transcriptional and post-transcriptional mechanisms (Figure 4).

PCAT1

PCAT1 was identified as an lncRNA overexpressed in a subset of prostate cancer patients, where it promotes cell proliferation (Prensner et al., 2011). PCAT1 expression is under negative control by the PRC2 complex. Consequently, patients could be stratified into groups with either PCAT1 lncRNA or PRC2 protein expression (Prensner et al., 2011). A SNP in a PCAT1 enhancer region (rs7463708) causes increased binding of the androgen receptor (AR) in complex with the ONECUT2 transcription factor, and consequently promotes elevated PCAT1

expression (Guo et al., 2016). PCAT1 lncRNA was shown to interact with an AR/LSD1 complex to promote expression of androgen-stimulated genes involved in prostate cancer progression (Guo et al., 2016; **Figure 4**). Meanwhile, PCAT1 has been implicated in many types of cancers, and further cancer-associated PCAT1 SNPs have been identified (Yang et al., 2019). The risk variant rs72725854 (A > T) for instance, locates to a prostate-specific enhancer and promotes expression of 8q24 lncRNAs PCAT1, PRNCR1 and PVT1 through SPDEF transcription factor recruitment (Walavalkar et al., 2020). Thus, similar to ANRIL and CCAT2, PCAT1 is a major oncogenic lncRNA and potential prognostic and therapeutic cancer marker.

HULC

HULC was first identified as an lncRNA overexpressed in hepatocellular carcinoma (Panzitt et al., 2007). Meanwhile, HULC lncRNA levels were shown to be altered in many types of cancer, through a complex network of transcriptional and post-transcriptional control mechanisms (Yu et al., 2017). The rs7763881 SNP in the HULC locus was associated with decreased risk of hepatocellular carcinoma development in patients with persistent HBV infection (Liu et al., 2012). Likewise, in colorectal and esophageal cancer studies, the rs7763881 SNP was found to be protective (Kang et al., 2015; Shaker et al., 2017). Several mechanisms have been proposed, through which HULC impacts mRNA expression and cancerogenesis. For instance, HULC was proposed to interact with YB-1 to promote phosphorylation and thereby dissociation of YB-1 from repressed messengers, including cyclin D1 and E1 mRNAs (Li et al., 2017a). Alternatively, HULC was suggested to bind to chromatin factors coREST and hnRNPk to promote the expression of

pro-inflammatory genes, such as IL6 (Chen et al., 2019b). In summary, while the association of lncRNA HULC with cancer is well-established, the precise mechanistic contribution of the rs7763881 SNP to HULC function remains to be determined.

CUPID1 and CUPID2

The 11q13 breast cancer risk locus contains several SNPs associated with estrogen-receptor-positive tumors. Betts et al. (2017) identified two lncRNAs, CUPID1 and CUPID2, which are transcribed from this locus in a head-to-head orientation. Transcription of both lncRNAs depends on the distal enhancer element PRE1. Importantly, the 11q13 breast cancer SNPs rs661204 and rs78540526 negatively affect the activatory PRE1 interaction with the CUPID1 promoter. Mechanistically, CUPID1 and CUPID2 locate to nuclear foci and promote DNA repair through homologous recombination but do not affect non-homologous end joining. This suggests that CUPID1 and CUPID2 are involved in the choice of the DNA repair pathway upon DNA double strand breaks (Betts et al., 2017). Thus, SNPs in the 11q13 locus likely contribute to breast cancer risk by disturbing PRE1 enhancer interaction with the shared promoter of DNA repair associated lncRNAs CUPID1 and CUPID2.

Besides ANRIL, CCAT2, PCAT1, HULC, and CUPID1/2, there are numerous other cancer-associated lncRNAs listed in the relevant databases (**Table 1**), several of which contain disease-associated SNPs (Ning et al., 2017; Bao et al., 2019). PTCSC3, for instance, is a tumor suppressor lncRNA located downstream of the rs944289 SNP at 14q.13.3, which predisposes to papillary thyroid carcinoma. The rs944289 SNP destroys a C/EBP α /

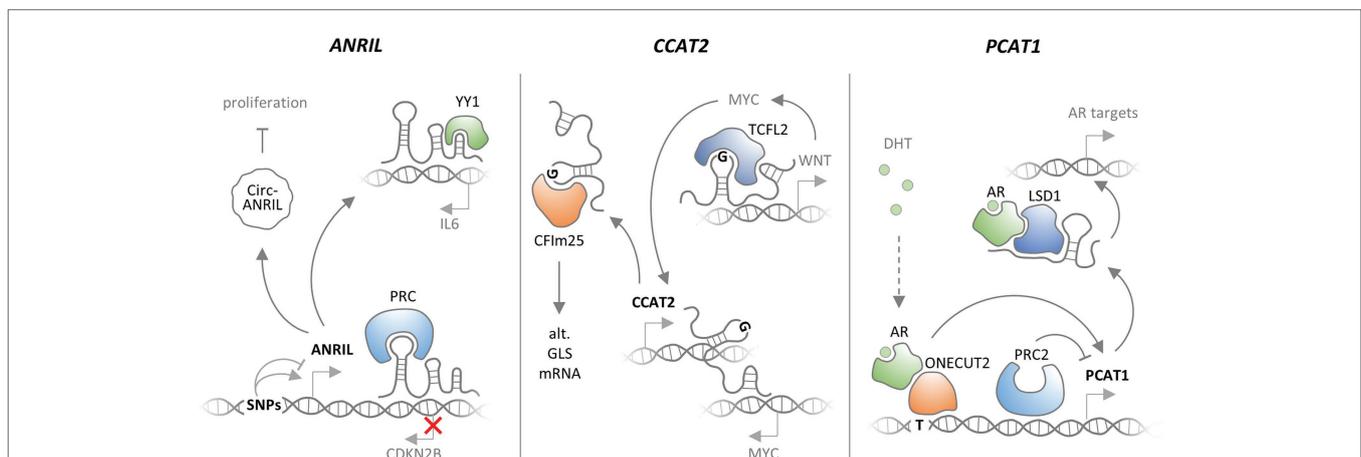


FIGURE 4 | Functional consequences of SNPs affecting lncRNAs. Different SNPs in the ANRIL lncRNA gene promoter promote and repress ANRIL expression, respectively. ANRIL associates with polycomb repressor complexes (PRCs) to silence CDKN2B expression. ANRIL may also bind the YY1 transcription factor to promote IL6 expression. Finally, a circular form of ANRIL was reported to promote proliferation. ANRIL promoter SNPs may impact on these cancer- and inflammation-relevant functions of ANRIL lncRNA. CCAT2 lncRNA is transcribed from chromosome 8 and may bind to the MYC locus on the same chromosome, thereby promoting MYC expression. Alternatively, CCAT2 may bind TCFL2 to promote WNT and thereby MYC expression. A SNP was shown to elevate CCAT2 expression. The SNP also promotes binding of CCAT2 to CFIm25, thereby promoting alternative splicing of GLS mRNA and thus cancer cell proliferation. LncRNA PCAT1 is under negative control by the PRC2 complex. A prostate cancer-associated SNP in the PCAT1 promoter confers increased binding of the dihydrotestosterone (DHT) associated androgen receptor (AR), in complex with the one-cut transcription factor, which leads to increased PCAT1 levels. In complex with AR and LSD1, PCAT1 promotes expression of cancer associated genes.

binding site, which reduces the expression of PTCSC3 (Jendrzewski et al., 2012). MIR2052HG is a breast-cancer relevant lncRNA, which harbors the rs4476990 and rs3802201 SNPs. Both SNPs foster estradiol-induced, ER α -dependent expression of MIR2052HG, which in turn promotes ER α expression and stability (Ingle et al., 2016). HOTAIR is not only a non-coding RNA regulator of HOX gene expression (Rinn et al., 2007), but also implicated in oncogenesis (Shah and Sukumar, 2010). Consequently, GWASs identified several cancer associated SNPs in the HOTAIR locus (Tang and Hann, 2018). One of these SNPs locates to an intronic enhancer and alters HOTAIR expression levels (Zhang et al., 2014). Further examples of lncRNAs with cancer SNPs include linc-PINT, MEG3, GAS5, PTENP1, H19, HOTTIP, CCAT1, or the P53-regulatory lncRNA network (Ning et al., 2017; Wang et al., 2017; Minotti et al., 2018; Carlevaro-Fita et al., 2020). Many other, less well-characterized lncRNA loci overlap with cancer risk variants or interact with distant risk variant regions through chromatin loops (Moradi Marjaneh et al., 2020). An emerging aspect of cancer-associated lncRNAs, with potential relevance to personalized therapy approaches, is their role in drug resistance. Some of the above-depicted lncRNAs with cancer SNPs, deletions and amplifications, such as HOTAIR, ANRIL, or MALAT1 were implicated in the development of treatment-resistant cancer cells (Liu et al., 2020). Taking into account, patients non-coding genome sequence information could therefore improve the determination of appropriate treatment regimens. Effective utilization of this information, however, demands a robust understanding of the causality between candidate lncRNA mutations and cancerogenesis or drug resistance. Recent research on lncRNA MALAT1 and MIR122HG associated mutations illustrates the difficulties regarding the establishment of such causality. Indel mutations in the MALAT1 and MIR122 host gene loci were identified in a pan-cancer analysis of somatic mutation hotspots (Rheinbay et al., 2020). Interestingly, MIR122 mutations were found outside of the miRNA encoded in this gene locus, and no alterations in miR-122 target-levels could be associated with this mutation, suggesting that they rather affect the miRNA host gene. Importantly, however, both MIR122 and MALAT1 indel mutations could not be confirmed as driver-mutations (Carlevaro-Fita et al., 2020; Rheinbay et al., 2020), and their roles in cancer thus remain unclear. It is therefore important, to aggregate information from diverse sources, including e.g., mutagenesis screens and mechanistic evidence, to establish causative relationships between lncRNA gene variations and cancer.

SNPs in Inflammation-Relevant lncRNA Loci

Besides cancer, infectious, and inflammatory diseases have remained a leading human health burden and a major cause of morbidity and mortality worldwide, despite great improvements in prevention and treatment (Holmes et al., 2017). Mortality associated with fulminant inflammation during sepsis for instance has remained difficult to control with broadly acting anti-inflammatory therapeutics, such as corticosteroids (Rochweg et al., 2018). An improved understanding of the molecular circuits driving immune-defense and inflammation is therefore

required to establish more targeted therapies for inflammation-induced pathologies. Recently, several lncRNAs were implicated in the cellular circuitries controlling inflammation, and genetic evidence for a role of lncRNAs in human inflammatory diseases has been provided (Castellanos-Rubio and Ghosh, 2019). This section depicts well-documented examples.

Lnc13

Celiac disease is a chronic illness, in which exposure to dietary gluten triggers inflammation in the intestine, often also accompanied by systemic symptoms (Ludvigsson et al., 2013; Tye-Din et al., 2018). Celiac disease affects genetically predisposed individuals, and risk-associated SNPs have been identified. The rs917997 SNP downstream of the IL18RAP locus for instance was reported to predispose to celiac disease (Hunt et al., 2008; Koskinen et al., 2009), the underlying mechanism however had remained unknown. In 2016, Castellanos-Rubio et al. (2016) reported an lncRNA, lnc13, which originates from the IL18RAP locus and harbors the rs917997 SNP. lnc13 represses inflammatory gene transcription by associating with hnRNP D in the nucleus and fostering promoter-binding of HDAC1 (Castellanos-Rubio et al., 2016). Expression levels of lnc13 were found to be reduced in celiac disease patients, which might contribute to increased inflammatory gene expression. Moreover, the rs917997 SNP in lnc13 was found to weaken the interaction of lnc13 with its interaction partner hnRNP D (Castellanos-Rubio et al., 2016). In summary, the celiac disease predisposing SNP rs917997 SNP seems to impact the function of an inflammation-limiting lncRNA by disturbing its association with nuclear protein machineries.

IFNG-AS1

The Tmevpg3 locus has long been known as a Theiler's virus susceptibility locus in mice, without a mechanistic explanation. Eventually, a lncRNA (IFNG-AS1, a.k.a. Tmevpg1 or NeST), mapping to the Tmevpg3 locus was shown to control expression of the nearby IFN γ gene, and a human homolog was identified (Vigneau et al., 2003). IFN γ is a major cytokine produced preferentially by T- and NK-cells and involved in Th1 responses and antibacterial defense. IFNG-AS is highly expressed in Th1-polarized T-cells, where it depends on the T-bet transcription factor and positively regulates IFN γ expression in cis (Collier et al., 2012; Padua et al., 2016; Petermann et al., 2019). The IFNG-AS gene overlaps with the IBD susceptibility SNP rs7134599 and was identified as the only out of several IBD SNP associated lncRNAs increased in ulcerative colitis patients (Mirza et al., 2015; Padua et al., 2016). Furthermore, elevated circulating IFNG-AS1 levels correlate with coronary artery disease severity (Xu and Shao, 2018). IFNG-AS1 was also shown to protect from severe systemic *Salmonella enterica* infection outcomes in mice, by promoting IFN γ expression (Gomez et al., 2013). Mechanistically, in the same study, IFNG-AS1 was shown to interact with the WDR5 histone-methyltransferase in the nucleus, to promote the deposition of activating histone methylation marks at the IFN γ locus. In summary, IFNG-AS1 is an IFN γ promoting chromatin-regulatory lncRNA, genetically associated with inflammation-linked human diseases.

LncRNA CCR5AS

The CCR5 chemokine receptor on the surface of T cells serves as a major HIV co-receptor during cell attachment and entry. Consequently, gene variants causing loss of CCR5 surface expression were found to confer resistance to HIV infection (McLaren and Carrington, 2015). GWASs have detected several SNPs, which impact on HIV infection (McLaren and Carrington, 2015). One of the top-ranking SNPs (rs1015164) on chromosome 3 maps down-stream of the CCRL2 locus, near an annotated antisense transcript (CCR5AS or RP11-24F11.2), which overlaps with the CCR5 gene body (McLaren et al., 2015; Kulkarni et al., 2019). The rs1015164 SNP marks an ATF1 binding site and enhances expression of CCR5AS. Knockdown of CCR5AS was found to reduce CCR5 mRNA abundance (Kulkarni et al., 2019). Mechanistically, CCR5AS was proposed to decoy RNA binding protein RALY, thereby enhancing CCR5 mRNA stability and consequently susceptibility to HIV infection (Kulkarni et al., 2019). Thus, lncRNA CCR5AS seems to constitute a central player in HIV resistance due to nucleotide variations in the CCR5 gene locus.

Besides the above depicted well-documented examples, further lncRNAs have been genetically linked to inflammation and infection, including e.g., streptococcal bacteremia associated lncRNA AC011288.2 (Kenyan Bacteraemia Study Group et al., 2016), but are still to be independently confirmed or explained mechanistically. Furthermore, lncRNAs with cancer-associated SNPs, such as NEAT1, Gas5, or ANRIL have been linked to inflammation (Castellanos-Rubio and Ghosh, 2019; Chen et al., 2019a). Additionally, increasing biochemical evidence suggests that many more lncRNAs participate in the cellular circuitries controlling inflammatory gene expression, including e.g., PACER, lincRNA-Cox2, or lincRNA-EPS (Chen et al., 2019a). Besides these established inflammation regulators, we recently determined dozens of so far uncharacterized lncRNAs, which are significantly up- or down-regulated upon bacterial infection or immune-activation of primary human cells (Aznaourova et al., 2020; Schulte et al., 2020). One of these lncRNAs, MaIL1, was identified by us as an essential component of the TLR4-TRIF immune-signaling pathway, mediating interferon expression in response to infectious agents. MaIL1 levels were increased in bronchoalveolar lavage fluid from patients with pulmonary infections (Aznaourova et al., 2020). Thus, increasing genetic and biochemical evidence suggests a vital participation of lncRNAs in human immunity, with potential relevance to patient stratification and prospective personalized therapies, as further discussed below (section Discussion).

SNPs in Other Disease-Associated lncRNA Loci

Many prevalent diseases involve an inflammatory component, including e.g., cardiovascular pathologies or cancer. Thus, it is not surprising, that several lncRNAs with disease-associated polymorphisms were implicated at the intersection of these three leading causes of death. Examples include NEAT1, Gas5, or ANRIL (see above), but also H19, LINC00305 and various

other lncRNAs (Giral et al., 2018; Minotti et al., 2018; Ramsuran et al., 2018; Castellanos-Rubio and Ghosh, 2019). Besides these three major human health burdens, lncRNA SNPs have also been associated with rather rare disorders, such as HELLP (hemolysis, elevated liver enzymes, and low platelets). LncRNA H19, LINC00305, and linc-HELLP are portrayed below as examples for lncRNAs linked to cardiovascular and rare diseases by GWAS.

H19

H19 was one of the first lncRNAs to be discovered and has been implicated in diseases, such as atherosclerosis (Han et al., 1996), though its non-coding nature was only recognized years after the H19 cDNA was first cloned. Originally, H19 was discovered as an RNA induced early during murine embryogenesis and muscle cell differentiation (Pachnis et al., 1988), which is reactivated during vascular smooth muscle injury (Kim et al., 1994). Meanwhile, several SNPs in the H19 lncRNA locus have been linked to increased risk of coronary artery disease (Gao et al., 2015) and predisposition to elevated blood-pressure (Tragante et al., 2014; Sung et al., 2018). Furthermore, H19 SNPs have been associated with several types of cancer (Verhaegh et al., 2008; Li et al., 2018a) and with elevated serum levels of H19 (Yang et al., 2015). Mechanistically, H19 was proposed to serve as a let-7 microRNA sponge in the cytosol (Kallen et al., 2013). On the other hand, H19 was reported to interact with chromatin factors in the nucleus to regulate gene expression (Monnier et al., 2013). Thus, while the role of lncRNA H19 in cardiovascular diseases is well-established, the precise molecular mode of H19 function, explaining the GWAS-identified disease predispositions is still under debate.

MIAT

MIAT is another lncRNA with well-documented roles in cardiovascular diseases and cancer. Originally identified as a myocardial infarction risk locus (Ishii et al., 2006), MIAT was soon suggested as a cardiovascular disease biomarker (Qu et al., 2017; Zhu et al., 2018). The rs2331291 (C > T) SNP in the fifth exon of MIAT was not only shown to be significantly associated with myocardial infarction, but also to promote MIAT1 expression, as well as its interaction with nuclear proteins (Ishii et al., 2006). The murine MIAT1 orthologue was found to interact with the nuclear matrix (Sone et al., 2007) and associate with splice factor SF1 *via* a conserved repeat motif, possibly to establish local nuclear splice-regulatory compartments (Tsuiji et al., 2011; Ishizuka et al., 2014). While the precise molecular contributions of MIAT to disease remain to be determined, the studies available so far suggest this lncRNA as a promising molecular prognosis marker among an array of further cardiovascular disease-relevant lncRNAs (Lozano-Vidal et al., 2019).

LINC00305

Besides H19 and MIAT, LINC00305 is another lncRNA with a potential role in cardiovascular diseases. A SNP in an intronic segment of the LINC00305 gene (rs2850711) had been associated

with atherosclerosis, which prompted Zhang et al. (2017b) to investigate the expression levels of this previously uncharacterized lncRNA in a patient cohort. LINC00305 expression was significantly increased in atherosclerotic plaques compared to normal artery samples (Zhang et al., 2017b). Furthermore, LINC00305 levels were higher in monocytes compared to endothelial and aortic smooth muscle cells. In monocytes, LINC00305 was found to promote inflammatory gene expression, probably involving NF κ B activation in a lipocalin-1 interacting membrane receptor (LIMR) and aryl-hydrocarbon receptor repressor (AHRR) dependent manner (Zhang et al., 2017b). In endothelial cells, LINC00305 was furthermore assigned a pro-apoptotic function (Zhang et al., 2017a). Increased expression of LINC00305 was also observed in serum of rheumatoid arthritis (RA) patients. Furthermore, RA patients carrying the rs2850711 polymorphism had significantly elevated LINC00305 and pro-inflammatory marker expression (Wahba et al., 2020). Together, these reports suggest that an intronic SNP in the LINC00305 gene predisposes to atherosclerosis and arthritis by increasing expression of apoptosis and inflammation promoting lncRNA LINC00305.

Linc-HELLP

HELLP (hemolysis, elevated liver enzymes, and low platelets) is a life-threatening pregnancy-associated syndrome, which entails initial placental dysfunction and eventual systemic maternal symptoms. An intergenic HELLP locus with several disease-associated SNPs on chromosome region 12q23.2 was identified, which gives rise to an lncRNA (linc-HELLP) and is flanked by the PMCH and IGF1 genes (van Dijk et al., 2012). Knockdown and RNA-Seq analysis of linc-HELLP in trophoblast cells suggested an involvement in cell survival and cell-cycle progression (van Dijk et al., 2012). Familial mutations in Linc-HELLP negatively affect trophoblast differentiation and alter the binding of protein interaction partners involved in splicing and ribosomal function (van Dijk et al., 2015). Thus, linc-HELLP constitutes another example of an lncRNA involved in critical cellular pathways and genetically associated with a severe human disease.

Besides the well-documented examples summarized in this review, many other disease-relevant lncRNA SNPs have been identified and are listed in the relevant databases (**Table 1**). In the context of cardiovascular diseases, for instance, SNPs affecting the well-studied lncRNA ANRIL were pointed out (Lozano-Vidal et al., 2019). Furthermore, aberrations in non-coding genome regions corresponding to transcribed enhancers potentially contribute to human diseases (Isoda et al., 2019). A closer investigation of mutations and polymorphisms affecting the transcription or function of non-coding RNAs generated from such regulatory elements might contribute additional insights into the roles of non-coding genome alterations in human diseases (see Discussion). The majority of disease-associated SNPs, however, are still awaiting to be independently reproduced or explained mechanistically. Furthermore, even for well-documented disease-associations of lncRNA SNPs, such as the

rs2331291 SNP in the MIAT locus, the aspect of causality between nucleotide variants and disease demands further investigation, as discussed below. Despite the still many open questions, however, a growing body of literature suggests that lncRNA gene loci, once regarded as evolutionary remnants and “junk DNA,” fulfill critical functions in human physiology and disease, and must be considered in next-gen diagnostics and personalized medicine approaches, similar to protein-coding genes.

DISCUSSION

Despite increasing evidence for an important role of lncRNAs in human diseases and their emerging application as biomarkers, their routine clinical implementation has not yet been achieved and only a fraction of the ~20,000 annotated lncRNAs has been studied to date. One reason might be that lncRNAs have not been included in large diagnostic studies for long, although disease-associated mutations affecting lncRNAs, such as ANRIL or IPW (see above) were known years before the release of comprehensive lncRNA annotations. Circumstances that might have delayed the inclusion of many lncRNAs in clinical cohort data analysis are the still provisional annotation status of many lncRNAs, their functional heterogeneity and the lack of approved pharmacological approaches to harness their therapeutic potential. While numerous recent publications have independently validated many annotated lncRNAs as non-coding transcripts with cell-type and condition-specific expression patterns, their molecular functions in diseases often remain unclear. Different from other classes of RNA, such as mRNAs, snRNAs, or miRNAs, lncRNAs do not seem to interact with a common protein-machinery (Aznaourova et al., 2020), which may render their investigation challenging. Recent studies have, however, substantially improved our understanding of lncRNA subcellular localization, coding-capacity, and interaction with protein-complexes (Carlevaro-Fita and Johnson, 2019; Aznaourova et al., 2020; Chen et al., 2020a), which will contribute to improved annotations and mechanistic models of lncRNA functions in disease. To promote the consideration of this lncRNA knowledge in clinical studies, better integration into medical NGS analysis workflows seems necessary. This demands not only the establishment of well-curated lncRNA knowledge databases but also of clinical informatics units as well as targeted training of young medical informatics staff. As a result of such measures, the advent of NGS in the clinics should increasingly accelerate our understanding of lncRNA functions and mutations in human diseases and eventually improve molecular diagnostics.

While research into the functions of lncRNAs in human diseases is still in its infancy, several lncRNAs have repeatedly been described as prognostic and risk stratification markers. Expression of PVT1 and FAL1, for instance, associates with advanced TNM stage in different types of cancer (Hu et al., 2014; Lu et al., 2017; Martini et al., 2017; Ni et al., 2017). Furthermore, prostate cancer patients could be stratified into distinct groups expressing either high levels of PRC 2 components or of the proliferation-promoting lncRNA PCAT1 (Prensner et al., 2011). Besides PCAT1, several other lncRNAs

have prognostic potential in prostate cancer (Arriaga-Canon et al., 2018). lncRNA PCA3 was even employed as a prostate cancer marker in a liquid biopsy test, though the test specificity remains debated (Morgan, 2019). Several of these lncRNAs contain disease-associated SNPs (see above), further suggesting their utility in molecular diagnostics. For example, it is conceivable that disease-associated lncRNA gene variants will be used in the future to forecast the success of cancer therapy. In fact, several lncRNAs affected by cancer-relevant SNPs or amplifications and deletions, including H19, PVT1, or ANRIL have been implicated in cancer drug resistance, as extensively reviewed elsewhere (Campos-Parra et al., 2018; Liu et al., 2020; Peng et al., 2020; Yuan et al., 2020). To solidify the clinical utility of these lncRNAs as diagnostic and prognostic markers for defined patient groups, larger cohorts and replication studies are desirable. In the case of lncRNA RMRP, such replication attempts could already pinpoint the occurrence of promoter mutations specifically to breast cancer patients (Rheinbay et al., 2017; Son et al., 2019). Other replication studies could causally link mutations affecting lncRNAs, such as RMRP, NEAT1, or LINC-PINT, to cancer (Rheinbay et al., 2017; Carlevaro-Fita et al., 2020). Such studies represent an important pillar of the current consolidation of medical lncRNA knowledge, which might eventually lead to improved molecular diagnostics and personalized therapies.

The increasing number of studies on disease-associated variations in lncRNA loci raises the question of potential advantages of lncRNAs as clinical biomarkers and targets, compared to established protein markers. Generally, lncRNAs seem to be more tissue and cell type specific than mRNAs (Cabili et al., 2011), which predestines them e.g., for use in liquid biopsy approaches. Effective clinical utilization of lncRNA knowledge thus demands the integration of data from sources, such as tissue expression catalogs and single-cell RNA-Seq studies of diseased and healthy tissue. Well-curated knowledge databases aggregating genetic and clinical study data, biochemical insights, and expression specificity information might pinpoint highly cell-type specific lncRNAs both as novel biomarkers and as prospective targets for precise pharmacological interventions in affected tissues and organ-systems. Another advantage of lncRNA- over protein-biomarkers is that with lncRNAs, functional biomolecules and their mutations can be directly quantified by sensitive, rapid, and cost-effective PCR- or NGS-based approaches, whereas mRNA levels, as intermediates, often do not reflect protein expression at the time-point of sample collection (Liu et al., 2016) and in the case of direct protein quantifications no information on mutations is provided. Irrespective of the potential advantages of lncRNAs over proteins as biomarkers, personalized medicine approaches based on whole genome profiling may only profit from a refined understanding of disease-associated driver and passenger mutations, taking into account the long overlooked mutations in lncRNA gene regions.

The continuous improvement of human transcriptome annotations has brought further classes of non-coding RNA into the focus of genetic studies. Besides small non-coding RNAs, such as snoRNAs (Deogharia and Majumder, 2018) or

microRNAs (Hrdlickova et al., 2014), genetic variation in non-coding RNAs generated from enhancer regions has received particular attention. Enhancers are typically marked by H3K4me1 signatures and by recruitment of co-activators of transcription, such as p300/CBP. Many enhancer regions in addition recruit RNAPolIII and generate eRNAs, which differ from GENCODE lncRNAs in that they are typically unspliced and non-polyadenylated (Kim et al., 2015). Besides eRNAs, RNAPolIII generates lncRNAs, transcribed from enhancer regions or recruited to super-enhancers, which are mostly tissue-specific enhancer clusters, bound to Mediator (Gardini and Shiekhhattar, 2015; Soibam, 2017). ncRNA transcription from such regions plays an important role in chromatin organization and enhancer-promoter communication. For instance, ncRNAs generated from enhancer regions can assist in the recruitment of factors such as CTCF or mediator, thereby contributing to chromatin loop formation (Gil and Ulitsky, 2020). Importantly, the act of transcription itself can exert important regulatory functions by affecting chromatin modification and accessibility (Ali and Grote, 2020). The act of transcription of non-coding RNA ThymoD, for instance, promotes CTCF binding site demethylation and thus formation of a chromatin-loop juxtaposing the Bcl11b enhancer and promoter, independent of the final RNA product (Isoda et al., 2017). In line with the role of Bcl11b in T cell commitment, disruption of ThymoD transcription in mice results in lymphoid malignancies (Isoda et al., 2017). In line with the critical roles of enhancers in the orchestration of gene expression in and outside of insulated chromatin domains [topologically associating domains (TADs); Fanucchi and Mhlanga, 2017], nucleotide variants and disruptions affecting enhancers, and potentially the transcription of eRNAs and lncRNAs generated from these regions, have been implicated in various diseases. This ranges from autoimmune diseases to mental disorders and cancer (Farh et al., 2015; Javierre et al., 2016; Teppo et al., 2016; Ren et al., 2017; Hauberg et al., 2019; Isoda et al., 2019; Lewis et al., 2019; Chen et al., 2020b; Yamagata et al., 2020). Thus, beyond lncRNAs encoded within coding gene regions or seemingly empty genomic space, genetic variation in transcribed regulatory DNA units, such as enhancers, contributes to human diseases. This indicates that the well-documented lncRNA gene polymorphisms and mutations summarized in the present review and in the relevant databases (Table 1) represent only a fraction of the aberrations in the non-coding genome contributing to disease.

As a result of the many recent reports on disease-relevant functions of human ncRNAs, their potential value as therapeutic targets is increasingly being recognized. mRNA-targeting antisense oligonucleotide drugs are arriving in the clinics (Shen and Corey, 2018), which makes targeting of lncRNAs seem feasible. Such RNA inhibitors are typically composed of synthetic nucleic acid mimics, such as locked nucleic acid (LNA), peptide nucleic acid (PNA), or phosphorodiamidate morpholino (PMO), which base-pair with RNA. Cellular delivery may be achieved through cell-penetrating peptides or packaging into lipid nanoparticles (Juliano, 2016; Nan and Zhang, 2018). The emerging clinical use of such RNA inhibitor technologies suggests that an improved understanding and extensive cataloguing of ncRNA

circuitries, altered during human diseases, may not only reveal valuable biomarkers and improve genome-wide diagnostics approaches but also open new therapeutic avenues.

Taken together, the advent of next-generation sequencing (NGS) technologies in clinical research and diagnostics has made it increasingly feasible to pinpoint non-coding gene variants and disease markers in a highly parallelized manner, and new chemical approaches allow to manipulate RNA levels therapeutically *in vivo*. Both will improve our understanding of the implications of the emerging class of lncRNAs in disease onset and progression. Currently, replication attempts with larger cohorts, and an improved understanding of the molecular functions of lncRNA markers seem necessary before the potential of lncRNAs in personalized medicine can be fully harnessed. However, continuous progress in the mechanistic understanding of lncRNAs in disease (DiStefano, 2018) and ambitious sequencing projects, such as the 100,000 Genomes Project (Sivapalaratnam and Bioresource, 2018) promise a further refinement of our knowledge about coding and non-coding disease-associated genomic alterations in the near future. Although it may still take years to achieve a comprehensive, clinically relevant understanding of the disease-contributions of lncRNAs,

whole-genome analysis of a patients coding and non-coding genome and transcriptome might eventually become part of the clinical routine, much like X-rays. Significant advances in molecular diagnostics, patient stratification, and our understanding of and intervention in molecular pathomechanisms are to be expected.

AUTHOR CONTRIBUTIONS

MA and NS screened, categorized, and summarized published literature. LS screened and categorized literature, wrote the manuscript draft, and prepared the figures. MA, NS, and BS proofread, corrected, and amended the manuscript draft. All authors contributed to the article and approved the submitted version.

FUNDING

PhD positions (MA and NS) were funded by the DFG SFB-TR84 “Innate Immunity of the Lung” (to LS and BS). The work was supported by the German Center for Lung Research (DZL).

REFERENCES

- Akrami, R., Jacobsen, A., Hoell, J., Schultz, N., Sander, C., and Larsson, E. (2013). Comprehensive analysis of long non-coding RNAs in ovarian cancer reveals global patterns and targeted DNA amplification. *PLoS One* 8:e80306. doi: 10.1371/journal.pone.0080306
- Ali, T., and Grote, P. (2020). Beyond the RNA-dependent function of lncRNA genes. *life* 9:e60583. doi: 10.7554/eLife.60583
- An, H., Williams, N. G., and Shelkova, T. A. (2018). NEAT1 and paraspeckles in neurodegenerative diseases: a missing lnc found? *Noncoding RNA Res.* 3, 243–252. doi: 10.1016/j.ncrna.2018.11.003
- Ang, C. E., Ma, Q., Wapinski, O. L., Fan, S., Flynn, R. A., Lee, Q. Y., et al. (2019). The novel lncRNA lnc-NR2F1 is pro-neurogenic and mutated in human neurodevelopmental disorders. *life* 8:e41770. doi: 10.7554/eLife.41770
- Antonescu, C. R., Agaram, N. P., Sung, Y. S., Zhang, L., Swanson, D., and Dickson, B. C. (2018). A distinct malignant epithelioid neoplasm with GLI1 gene rearrangements, frequent S100 protein expression, and metastatic potential: expanding the spectrum of pathologic entities with ACTB/MALAT1/PTCH1-GLI1 fusions. *Am. J. Surg. Pathol.* 42, 553–560. doi: 10.1097/PAS.0000000000001010
- Aplan, P. D. (2006). Causes of oncogenic chromosomal translocation. *Trends Genet.* 22, 46–55. doi: 10.1016/j.tig.2005.10.002
- Arriaga-Canon, C., De La Rosa-Velazquez, I. A., Gonzalez-Barríos, R., Montiel-Manriquez, R., Oliva-Rico, D., Jimenez-Trejo, F., et al. (2018). The use of long non-coding RNAs as prognostic biomarkers and therapeutic targets in prostate cancer. *Oncotarget* 9, 20872–20890. doi: 10.18632/oncotarget.25038
- Athie, A., Marchese, F. P., Gonzalez, J., Lozano, T., Raimondi, I., Juvvuna, P. K., et al. (2020). Analysis of copy number alterations reveals the lncRNA ALAL-1 as a regulator of lung cancer immune evasion. *J. Cell Biol.* 219:e201908078. doi: 10.1083/jcb.201908078
- Atianand, M. K., Caffrey, D. R., and Fitzgerald, K. A. (2017). Immunobiology of long noncoding RNAs. *Annu. Rev. Immunol.* 35, 177–198. doi: 10.1146/annurev-immunol-041015-055459
- Aznaourova, M., Janga, H., Sefried, S., Kaufmann, A., Dorna, J., Volkens, S. M., et al. (2020). Noncoding RNA MaLL1 is an integral component of the TLR4-TRIF pathway. *Proc. Natl. Acad. Sci. U. S. A.* 117, 9042–9053. doi: 10.1073/pnas.1920393117
- Bahuau, M., Vidaud, D., Jenkins, R. B., Bieche, I., Kimmel, D. W., Assouline, B., et al. (1998). Germ-line deletion involving the INK4 locus in familial proneness to melanoma and nervous system tumors. *Cancer Res.* 58, 2298–2303
- Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., and Dong, D. (2019). lncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.* 47, D1034–D1037. doi: 10.1093/nar/gky905
- Bei, J. X., Li, Y., Jia, W. H., Feng, B. J., Zhou, G., Chen, L. Z., et al. (2010). A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat. Genet.* 42, 599–603. doi: 10.1038/ng.601
- Betts, J. A., Moradi Marjaneh, M., Al-Ejeh, F., Lim, Y. C., Shi, W., Sivakumaran, H., et al. (2017). Long noncoding RNAs CUPID1 and CUPID2 mediate breast cancer risk at 11q13 by modulating the response to DNA damage. *Am. J. Hum. Genet.* 101, 255–266. doi: 10.1016/j.ajhg.2017.07.007
- Blackwood, D. H., Fordyce, A., Walker, M. T., St Clair, D. M., Porteous, D. J., and Muir, W. J. (2001). Schizophrenia and affective disorders—co-segregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. *Am. J. Hum. Genet.* 69, 428–433. doi: 10.1086/321969
- Bodega, B., Ramirez, G. D., Grasser, F., Cheli, S., Brunelli, S., Mora, M., et al. (2009). Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol.* 7:41. doi: 10.1186/1741-7007-7-41
- Bodmer, D., Eleveld, M., Kater-Baats, E., Janssen, I., Janssen, B., Weterman, M., et al. (2002). Disruption of a novel MFS transporter gene, DIRC2, by a familial renal cell carcinoma-associated t(2;3)(q35;q21). *Hum. Mol. Genet.* 11, 641–649. doi: 10.1093/hmg/11.6.641
- Bodmer, D., Schepens, M., Eleveld, M. J., Schoenmakers, E. F., and Geurts van Kessel, A. (2003). Disruption of a novel gene, DIRC3, and expression of DIRC3-HSPBAP1 fusion transcripts in a case of familial renal cell cancer and t(2;3)(q35;q21). *Genes Chromosom. Cancer* 38, 107–116. doi: 10.1002/gcc.10243
- Boloix, A., Masanas, M., Jimenez, C., Antonelli, R., Soriano, A., Roma, J., et al. (2019). Long non-coding RNA PVT1 as a prognostic and therapeutic target in pediatric cancer. *Front. Oncol.* 9:1173. doi: 10.3389/fonc.2019.01173
- Broadbent, H. M., Peden, J. F., Lorkowski, S., Goel, A., Ongen, H., Green, F., et al. (2008). Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet.* 17, 806–814. doi: 10.1093/hmg/ddm352

- Burd, C. E., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z., and Sharpless, N. E. (2010). Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.* 6:e1001233. doi: 10.1371/journal.pgen.1001233
- Butler, M. G., Miller, J. L., and Forster, J. L. (2019). Prader-Willi syndrome—clinical genetics, diagnosis and treatment approaches: an update. *Curr. Pediatr. Rev.* 15, 207–244. doi: 10.2174/1573396315666190716120925
- Cabianca, D. S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., et al. (2012). A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* 149, 819–831. doi: 10.1016/j.cell.2012.03.035
- Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. doi: 10.1101/gad.17446611
- Campos-Parra, A. D., Lopez-Urrutia, E., Orozco Moreno, L. T., Lopez-Camarillo, C., Meza-Menchaca, T., Figueroa Gonzalez, G., et al. (2018). Long non-coding RNAs as new master regulators of resistance to systemic treatments in breast cancer. *Int. J. Mol. Sci.* 19:2711. doi: 10.3390/ijms19092711
- Carlevaro-Fita, J., and Johnson, R. (2019). Global positioning system: understanding long noncoding RNAs through subcellular localization. *Mol. Cell* 73, 869–883. doi: 10.1016/j.molcel.2019.02.008
- Carlevaro-Fita, J., Lanzos, A., Feuerbach, L., Hong, C., Mas-Ponte, D., Pedersen, J. S., et al. (2020). Cancer lncRNA census reveals evidence for deep functional conservation of long noncoding RNAs in tumorigenesis. *Commun. Biol.* 3:56. doi: 10.1038/s42003-019-0741-7
- Castellanos-Rubio, A., Fernandez-Jimenez, N., Kratchmarov, R., Luo, X., Bhagat, G., Green, P. H., et al. (2016). A long noncoding RNA associated with susceptibility to celiac disease. *Science* 352, 91–95. doi: 10.1126/science.aad0467
- Castellanos-Rubio, A., and Ghosh, S. (2019). Disease-associated SNPs in inflammation-related lncRNAs. *Front. Immunol.* 10:420. doi: 10.3389/fimmu.2019.00420
- Chen, J., Ao, L., and Yang, J. (2019a). Long non-coding RNAs in diseases related to inflammation and immunity. *Ann. Transl. Med.* 7:494. doi: 10.21037/atm.2019.08.37
- Chen, J., Brunner, A. D., Cogan, J. Z., Nunez, J. K., Fields, A. P., Adamson, B., et al. (2020a). Pervasive functional translation of noncanonical human open reading frames. *Science* 367, 1140–1146. doi: 10.1126/science.aay0262
- Chen, Y., Fu, Y., Song, Y. F., and Li, N. (2019b). Increased expression of lncRNA UCA1 and HULC is required for pro-inflammatory response during LPS induced sepsis in endothelial cells. *Front. Physiol.* 10:608. doi: 10.3389/fphys.2019.00608
- Chen, X. F., Guo, M. R., Duan, Y. Y., Jiang, F., Wu, H., Dong, S. S., et al. (2020b). Multiomics dissection of molecular regulatory mechanisms underlying autoimmune-associated noncoding SNPs. *JCI Insight* 5:e136477. doi: 10.1172/jci.insight.136477
- Chen, X., Hao, Y., Cui, Y., Fan, Z., He, S., Luo, J., et al. (2017). LncVar: a database of genetic variation associated with long non-coding genes. *Bioinformatics* 33, 112–118. doi: 10.1093/bioinformatics/btw581
- Chubb, J. E., Bradshaw, N. J., Soares, D. C., Porteous, D. J., and Millar, J. K. (2008). The DISC locus in psychiatric illness. *Mol. Psychiatry* 13, 36–64. doi: 10.1038/sj.mp.4002106
- Coe, E. A., Tan, J. Y., Shapiro, M., Louphrasitthiphol, P., Bassett, A. R., Marques, A. C., et al. (2019). The MITF-SOX10 regulated long non-coding RNA DIRC3 is a melanoma tumour suppressor. *PLoS Genet.* 15:e1008501. doi: 10.1371/journal.pgen.1008501
- Collier, S. P., Collins, P. L., Williams, C. L., Boothby, M. R., and Aune, T. M. (2012). Cutting edge: influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *J. Immunol.* 189, 2084–2088. doi: 10.4049/jimmunol.1200774
- Cunnington, M. S., Santibanez Koref, M., Mayosi, B. M., Burn, J., and Keavney, B. (2010). Chromosome 9p21 SNPs associated with multiple disease phenotypes correlate with ANRIL expression. *PLoS Genet.* 6:e1000899. doi: 10.1371/journal.pgen.1000899
- Delihas, N. (2018). A family of long intergenic non-coding RNA genes in human chromosomal region 22q11.2 carry a DNA translocation breakpoint/AT-rich sequence. *PLoS One* 13:e0195702. doi: 10.1371/journal.pone.0195702
- Deogharia, M., and Majumder, M. (2018). Guide snoRNAs: drivers or passengers in human disease? *Biology* 8:1. doi: 10.3390/biology8010001
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111
- Di Donato, N., Riess, A., Hackmann, K., Rump, A., Huebner, A., von der Hagen, M., et al. (2012). Macrocephaly, obesity, mental (intellectual) disability, and ocular abnormalities: alternative definition and further delineation of MOMO syndrome. *Am. J. Med. Genet.* 158A, 2857–2862. doi: 10.1002/ajmg.a.35481
- Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical, Saxena, R., Voight, B. F., Lyssenko, V., Burt, N. P., de Bakker, P. I., et al. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316, 1331–1336. doi: 10.1126/science.1142358
- DiStefano, J. K. (2018). The emerging role of long noncoding RNAs in human disease. *Methods Mol. Biol.* 1706, 91–110. doi: 10.1007/978-1-4939-7471-9_6
- Du, Z., Fei, T., Verhaak, R. G., Su, Z., Zhang, Y., Brown, M., et al. (2013). Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat. Struct. Mol. Biol.* 20, 908–913. doi: 10.1038/nsmb.2591
- Eissmann, M., Gutschner, T., Hammerle, M., Gunther, S., Caudron-Herger, M., Gross, M., et al. (2012). Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biol.* 9, 1076–1087. doi: 10.4161/rna.21089
- Fang, S., Zhang, L., Guo, J., Niu, Y., Wu, Y., Li, H., et al. (2018). NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res.* 46, D308–D314. doi: 10.1093/nar/gkx1107
- Fanucchi, S., and Mhlanga, M. M. (2017). Enhancer-derived lncRNAs regulate genome architecture: fact or fiction? *Trends Genet.* 33, 375–377. doi: 10.1016/j.tig.2017.03.004
- Farh, K. K. H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W. J., Beik, S., et al. (2015). Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 518, 337–343. doi: 10.1038/nature13835
- Fujimoto, A., Furuta, M., Totoki, Y., Tsunoda, T., Kato, M., Shiraiishi, Y., et al. (2016). Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat. Genet.* 48, 500–509. doi: 10.1038/ng.3547
- Fulco, C. P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S. R., Perez, E. M., et al. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* 354, 769–773. doi: 10.1126/science.aag2445
- Gabay, M., Li, Y., and Felsher, D. W. (2014). MYC activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harb. Pers. Med.* 4:a014241. doi: 10.1101/cshperspect.a014241
- Gao, Y., Wang, P., Wang, Y., Ma, X., Zhi, H., Zhou, D., et al. (2019). Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers. *Nucleic Acids Res.* 47, D1028–D1033. doi: 10.1093/nar/gky1096
- Gao, W., Zhu, M., Wang, H., Zhao, S., Zhao, D., Yang, Y., et al. (2015). Association of polymorphisms in long non-coding RNA H19 with coronary artery disease risk in a Chinese population. *Mutat. Res.* 772, 15–22. doi: 10.1016/j.mrfmmm.2014.12.009
- Gardini, A., and Shiekhata, R. (2015). The many faces of long noncoding RNAs. *FEBS J.* 282, 1647–1657. doi: 10.1111/febs.13101
- Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* 21, 102–117. doi: 10.1038/s41576-019-0184-5
- Giral, H., Landmesser, U., and Kratzer, A. (2018). Into the wild: GWAS exploration of non-coding RNAs. *Front. Cardiovasc. Med.* 5:181. doi: 10.3389/fcvm.2018.00181
- Gomez, J. A., Wapinski, O. L., Yang, Y. W., Bureau, J. F., Gopinath, S., Monack, D. M., et al. (2013). The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. *Cell* 152, 743–754. doi: 10.1016/j.cell.2013.01.015
- Graham, R. P., Nair, A. A., Davila, J. I., Jin, L., Jen, J., Sukov, W. R., et al. (2017). Gastroblastoma harbors a recurrent somatic MALAT1-GLI1 fusion gene. *Mod. Pathol.* 30, 1443–1452. doi: 10.1038/modpathol.2017.68
- Guo, H., Ahmed, M., Zhang, F., Yao, C. Q., Li, S., Liang, Y., et al. (2016). Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. *Nat. Genet.* 48, 1142–1150. doi: 10.1038/ng.3637

- Han, D. K., Khaing, Z. Z., Pollock, R. A., Haudenschild, C. C., and Liao, G. (1996). H19, a marker of developmental transition, is reexpressed in human atherosclerotic plaques and is regulated by the insulin family of growth factors in cultured rabbit smooth muscle cells. *J. Clin. Invest.* 97, 1276–1285. doi: 10.1172/JCI118543
- Han, C., Sun, L. Y., Wang, W. T., Sun, Y. M., and Chen, Y. Q. (2019). Non-coding RNAs in cancers with chromosomal rearrangements: the signatures, causes, functions and implications. *J. Mol. Cell Biol.* 11, 886–898. doi: 10.1093/jmcb/mjz080
- Harismendy, O., Notani, D., Song, X., Rahim, N. G., Tanasa, B., Heintzman, N., et al. (2011). 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature* 470, 264–268. doi: 10.1038/nature09753
- Haugberg, M. E., Fullard, J. F., Zhu, L., Cohain, A. T., Giambartolomei, C., Misir, R., et al. (2019). Differential activity of transcribed enhancers in the prefrontal cortex of 537 cases with schizophrenia and controls. *Mol. Psychiatry* 24, 1685–1695. doi: 10.1038/s41380-018-0059-8
- He, Y., Jing, Y., Wei, F., Tang, Y., Yang, L., Luo, J., et al. (2018). Long non-coding RNA PVT1 predicts poor prognosis and induces radioresistance by regulating DNA repair and cell apoptosis in nasopharyngeal carcinoma. *Cell Death Dis.* 9:235. doi: 10.1038/s41419-018-0265-y
- He, X., Ou, C., Xiao, Y., Han, Q., Li, H., and Zhou, S. (2017). LncRNAs: key players and novel insights into diabetes mellitus. *Oncotarget* 8, 71325–71341. doi: 10.18632/oncotarget.19921
- Helgadottir, A., Thorleifsson, G., Manolescu, A., Gretarsdottir, S., Blondal, T., Jonasdottir, A., et al. (2007). A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 316, 1491–1493. doi: 10.1126/science.1142842
- Hindorf, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., et al. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9362–9367. doi: 10.1073/pnas.0903103106
- Holdt, L. M., Stahring, A., Sass, K., Pichler, G., Kulak, N. A., Wilfert, W., et al. (2016). Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat. Commun.* 7:12429. doi: 10.1038/ncomms12429
- Holmes, K. K., Bertozzi, S., Bloom, B. R., Jha, P., Gelband, H., DeMaria, L. M., et al. (2017). “Major infectious diseases: key messages from disease control priorities” in *Major infectious diseases. 3rd Edn.* Washington, DC: The International Bank for Reconstruction and Development/The World Bank.
- Hon, C. C., Ramilowski, J. A., Harshbarger, J., Bertin, N., Rackham, O. J., Gough, J., et al. (2017). An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* 543, 199–204. doi: 10.1038/nature21374
- Hrdlickova, B., de Almeida, R. C., Borek, Z., and Withoff, S. (2014). Genetic variation in the non-coding genome: involvement of micro-RNAs and long non-coding RNAs in disease. *Biochim. Biophys. Acta* 1842, 1910–1922. doi: 10.1016/j.bbdis.2014.03.011
- Hu, X., Feng, Y., Zhang, D., Zhao, S. D., Hu, Z., Greshock, J., et al. (2014). A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. *Cancer Cell* 26, 344–357. doi: 10.1016/j.ccr.2014.07.009
- Huarte, M. (2015). The emerging role of lncRNAs in cancer. *Nat. Med.* 21, 1253–1261. doi: 10.1038/nm.3981
- Hunt, K. A., Zhernakova, A., Turner, G., Heap, G. A., Franke, L., Bruinenberg, M., et al. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.* 40, 395–402. doi: 10.1038/ng.102
- Ingle, J. N., Xie, F., Ellis, M. J., Goss, P. E., Shepherd, L. E., Chapman, J. W., et al. (2016). Genetic polymorphisms in the long noncoding RNA MIR2052HG offer a pharmacogenomic basis for the response of breast cancer patients to aromatase inhibitor therapy. *Cancer Res.* 76, 7012–7023. doi: 10.1158/0008-5472.CAN-16-1371
- Ishii, N., Ozaki, K., Sato, H., Mizuno, H., Susumu, S., Takahashi, A., et al. (2006). Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J. Hum. Genet.* 51, 1087–1099. doi: 10.1007/s10038-006-0070-9
- Ishizuka, A., Hasegawa, Y., Ishida, K., Yanaka, K., and Nakagawa, S. (2014). Formation of nuclear bodies by the lncRNA Gomafu-associating proteins Celf3 and SF1. *Genes Cells* 19, 704–721. doi: 10.1111/gtc.12169
- Isoda, T., Moore, A. J., He, Z., Chandra, V., Aida, M., Denholtz, M., et al. (2017). Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* 171, 103.e118–119.e118. doi: 10.1016/j.cell.2017.09.001
- Isoda, T., Morio, T., and Takagi, M. (2019). Noncoding RNA transcription at enhancers and genome folding in cancer. *Cancer Sci.* 110, 2328–2336. doi: 10.1111/cas.14107
- Janga, H., Aznaourova, M., Boldt, F., Damm, K., Grunweller, A., and Schulte, L. N. (2018). Cas9-mediated excision of proximal DNaseI/H3K4me3 signatures confers robust silencing of microRNA and long non-coding RNA genes. *PLoS One* 13:e0193066. doi: 10.1371/journal.pone.0193066
- Javierre, B. M., Burren, O. S., Wilder, S. P., Kreuzhuber, R., Hill, S. M., Sewitz, S., et al. (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* 167:1369. doi: 10.1016/j.cell.2016.09.037
- Jendrzewski, J., He, H., Radomska, H. S., Li, W., Tomsic, J., Liyanarachchi, S., et al. (2012). The polymorphism rs944289 predisposes to papillary thyroid carcinoma through a large intergenic noncoding RNA gene of tumor suppressor type. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8646–8651. doi: 10.1073/pnas.1205654109
- Ji, J., Dai, X., Yeung, S. J., and He, X. (2019). The role of long non-coding RNA GAS₅ in cancers. *Cancer Manag. Res.* 11, 2729–2737. doi: 10.2147/CMAR.S189052
- Jin, K., Wang, S., Zhang, Y., Xia, M., Mo, Y., Li, X., et al. (2019). Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis. *Cell. Mol. Life Sci.* 76, 4275–4289. doi: 10.1007/s00018-019-03222-1
- Juliano, R. L. (2016). The delivery of therapeutic oligonucleotides. *Nucleic Acids Res.* 44, 6518–6548. doi: 10.1093/nar/gkw236
- Kacprzak, M., Chrzanoska, M., Skoczylas, B., Moczulska, H., Borowiec, M., and Sieroszewski, P. (2016). Genetic causes of recurrent miscarriages. *Ginek. Pol.* 87, 722–726. doi: 10.5603/GP.2016.0075
- Kallen, A. N., Zhou, X. B., Xu, J., Qiao, C., Ma, J., Yan, L., et al. (2013). The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol. Cell* 52, 101–112. doi: 10.1016/j.molcel.2013.08.027
- Kang, M., Sang, Y., Gu, H., Zheng, L., Wang, L., Liu, C., et al. (2015). Long noncoding RNAs POLR2E rs3787016 C/T and HULC rs7763881 A/C polymorphisms are associated with decreased risk of esophageal cancer. *Tumour Biol.* 36, 6401–6408. doi: 10.1007/s13277-015-3328-z
- Kenyan Bacteraemia Study Group, Wellcome Trust Case Control Consortium, Rautanen, A., Pirinen, M., Mills, T. C., Rockett, K. A., et al. (2016). Polymorphism in a lincRNA associates with a doubled risk of pneumococcal bacteremia in Kenyan children. *Am. J. Hum. Genet.* 98, 1092–1100. doi: 10.1016/j.ajhg.2016.03.025
- Kim, T. K., Hemberg, M., and Gray, J. M. (2015). Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. *Cold Spring Harb. Perspect. Biol.* 7:a018622. doi: 10.1101/cshperspect.a018622
- Kim, D. K., Zhang, L., Dzau, V. J., and Pratt, R. E. (1994). H19, a developmentally regulated gene, is reexpressed in rat vascular smooth muscle cells after injury. *J. Clin. Invest.* 93, 355–360. doi: 10.1172/JCI116967
- Kino, T., Hurt, D. E., Ichijo, T., Nader, N., and Chrousos, G. P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.* 3:ra8. doi: 10.1126/scisignal.2000568
- Klopocki, E., Hennig, B. P., Dathe, K., Koll, R., de Ravel, T., Baten, E., et al. (2010). Deletion and point mutations of PTHLH cause brachydactyly type E. *Am. J. Hum. Genet.* 86, 434–439. doi: 10.1016/j.ajhg.2010.01.023
- Koob, M. D., Moseley, M. L., Schut, L. J., Benzow, K. A., Bird, T. D., Day, J. W., et al. (1999). An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat. Genet.* 21, 379–384. doi: 10.1038/7710
- Koskinen, L. L., Einarsdottir, E., Dukes, E., Heap, G. A., Dubois, P., Korponay-Szabo, I. R., et al. (2009). Association study of the IL18RAP locus in three European populations with coeliac disease. *Hum. Mol. Genet.* 18, 1148–1155. doi: 10.1093/hmg/ddn438
- Kulkarni, S., Lied, A., Kulkarni, V., Rucevic, M., Martin, M. P., Walker-Sperling, V., et al. (2019). CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. *Nat. Immunol.* 20, 824–834. doi: 10.1038/s41590-019-0406-1
- Lazar, D. C., Morris, K. V., and Saayman, S. M. (2016). The emerging role of long non-coding RNAs in HIV infection. *Virus Res.* 212, 114–126. doi: 10.1016/j.virusres.2015.07.023

- Leucci, E., Vendramin, R., Spinazzi, M., Laurette, P., Fiers, M., Wouters, J., et al. (2016). Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* 531, 518–522. doi: 10.1038/nature17161
- Lewis, M. W., Li, S., and Franco, H. L. (2019). Transcriptional control by enhancers and enhancer RNAs. *Transcription* 10, 171–186. doi: 10.1080/21541264.2019.1695492
- Li, Z., Dou, P., Liu, T., and He, S. (2017b). Application of long noncoding rnas in osteosarcoma: biomarkers and therapeutic targets. *Cell. Physiol. Biochem.* 42, 1407–1419. doi: 10.1159/000479205
- Li, L., Guo, G., Zhang, H., Zhou, B., Bai, L., Chen, H., et al. (2018a). Association between H19 SNP rs217727 and lung cancer risk in a Chinese population: a case control study. *BMC Med. Genet.* 19:136. doi: 10.1186/s12881-018-0573-1
- Li, S., Li, J., Chen, C., Zhang, R., and Wang, K. (2018b). Pan-cancer analysis of long non-coding RNA NEAT1 in various cancers. *Genes Dis.* 5, 27–35. doi: 10.1016/j.gendis.2017.11.003
- Li, D., Liu, X., Zhou, J., Hu, J., Zhang, D., Liu, J., et al. (2017a). Long noncoding RNA HULC modulates the phosphorylation of YB-1 through serving as a scaffold of extracellular signal-regulated kinase and YB-1 to enhance hepatocarcinogenesis. *Hepatology* 65, 1612–1627. doi: 10.1002/hep.29010
- Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R. S., et al. (2013). CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. *Genome Res.* 23, 1446–1461. doi: 10.1101/gr.152942.112
- Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell* 165, 535–550. doi: 10.1016/j.cell.2016.03.014
- Liu, K., Gao, L., Ma, X., Huang, J. J., Chen, J., Zeng, L., et al. (2020). Long non-coding RNAs regulate drug resistance in cancer. *Mol. Cancer* 19:54. doi: 10.1186/s12943-020-01162-0
- Liu, Y., Pan, S., Liu, L., Zhai, X., Liu, J., Wen, J., et al. (2012). A genetic variant in long non-coding RNA HULC contributes to risk of HBV-related hepatocellular carcinoma in a Chinese population. *PLoS One* 7:e35145. doi: 10.1371/journal.pone.0053178
- Lozano-Vidal, N., Bink, D. I., and Boon, R. A. (2019). Long noncoding RNA in cardiac aging and disease. *J. Mol. Cell Biol.* 11, 860–867. doi: 10.1093/jmcb/mjz046
- Lu, D., Luo, P., Wang, Q., Ye, Y., and Wang, B. (2017). lncRNA PVT1 in cancer: a review and meta-analysis. *Clin. Chim. Acta* 474, 1–7. doi: 10.1016/j.cca.2017.08.038
- Lu, Z., Pannunzio, N. R., Greisman, H. A., Casero, D., Parekh, C., and Lieber, M. R. (2015). Convergent BCL6 and lncRNA promoters demarcate the major breakpoint region for BCL6 translocations. *Blood* 126, 1730–1731. doi: 10.1182/blood-2015-07-657999
- Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62, 43–52. doi: 10.1136/gutjnl-2011-301346
- Luo, Z., Xiao, L., Li, J., Dong, B., and Wang, C. (2019). Integrative analysis reveals driver long non-coding RNAs in osteosarcoma. *Medicine* 98:e14302. doi: 10.1097/MD.00000000000018446
- Lv, X., Li, Y., Li, Y., Li, H., Zhou, L., Wang, B., et al. (2019). FAL1: a critical oncogenic long non-coding RNA in human cancers. *Life Sci.* 236:116918. doi: 10.1016/j.lfs.2019.116918
- Ma, L., Cao, J., Liu, L., Du, Q., Li, Z., Zou, D., et al. (2019). LncBook: a curated knowledgebase of human long non-coding RNAs. *Nucleic Acids Res.* 47:2699. doi: 10.1093/nar/gkz073
- Ma, Z., Gu, S., Song, M., Yan, C., Hui, B., Ji, H., et al. (2017). Long non-coding RNA SNHG17 is an unfavourable prognostic factor and promotes cell proliferation by epigenetically silencing P57 in colorectal cancer. *Mol. Biosyst.* 13, 2350–2361. doi: 10.1039/C7MB00280G
- Ma, L., Li, A., Zou, D., Xu, X., Xia, L., Yu, J., et al. (2015). LncRNAWiki: harnessing community knowledge in collaborative curation of human long non-coding RNAs. *Nucleic Acids Res.* 43, D187–D192. doi: 10.1093/nar/gku1167
- Maass, P. G., Rump, A., Schulz, H., Stricker, S., Schulze, L., Platzer, K., et al. (2012). A misplaced lncRNA causes brachydactyly in humans. *J. Clin. Investig.* 122, 3990–4002. doi: 10.1172/JCI65508
- MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* 45, D896–D901. doi: 10.1093/nar/gkw1133
- Martin, A. N., and Li, Y. (2007). RNase MRP RNA and human genetic diseases. *Cell Res.* 17, 219–226. doi: 10.1038/sj.cr.7310120
- Martini, P., Paracchini, L., Caratti, G., Mello-Grand, M., Fruscio, R., Beltrame, L., et al. (2017). lncRNAs as novel indicators of patients' prognosis in stage I epithelial ovarian cancer: a retrospective and multicentric study. *Clin. Clin. Cancer Res.* 23, 2356–2366. doi: 10.1158/1078-0432.CCR-16-1402
- Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., et al. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190–1195. doi: 10.1126/science.1222794
- McLaren, P. J., and Carrington, M. (2015). The impact of host genetic variation on infection with HIV-1. *Nat. Immunol.* 16, 577–583. doi: 10.1038/ni.3147
- McLaren, P. J., Coulonges, C., Bartha, I., Lenz, T. L., Deutsch, A. J., Bashirova, A., et al. (2015). Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. *Proc. Natl. Acad. Sci. U. S. A.* 112, 14658–14663. doi: 10.1073/pnas.1514867112
- McPherson, R., Pertsemliadis, A., Kavaslar, N., Stewart, A., Roberts, R., Cox, D. R., et al. (2007). A common allele on chromosome 9 associated with coronary heart disease. *Science* 316, 1488–1491. doi: 10.1126/science.1142447
- Millar, J. K., Wilson-Annan, J. C., Anderson, S., Christie, S., Taylor, M. S., Semple, C. A., et al. (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* 9, 1415–1423. doi: 10.1093/hmg/9.9.1415
- Minotti, L., Agnoletto, C., Baldassari, F., Corra, F., and Volinia, S. (2018). SNPs and somatic mutation on long non-coding RNA: new frontier in the cancer studies? *High Throughput* 7:34. doi: 10.3390/ht7040034
- Mirza, A. H., Berthelsen, C. H., Seemann, S. E., Pan, X., Frederiksen, K. S., Vilien, M., et al. (2015). Transcriptomic landscape of lncRNAs in inflammatory bowel disease. *Genome Med.* 7:39. doi: 10.1186/s13073-015-0162-2
- Monnier, P., Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S., and Dandolo, L. (2013). H19 lncRNA controls gene expression of the imprinted gene network by recruiting MBD1. *Proc. Natl. Acad. Sci. U. S. A.* 110, 20693–20698. doi: 10.1073/pnas.1310201110
- Moradi Marjaneh, M., Beesley, J., O'Mara, T. A., Mukhopadhyay, P., Koufariotis, L. T., Kazakoff, S., et al. (2020). Non-coding RNAs underlie genetic predisposition to breast cancer. *Genome Biol.* 21:7. doi: 10.1186/s13059-019-1876-z
- Moretti-Ferreira, D., Koiffmann, C. P., Listik, M., Setian, N., and Wajntal, A. (1993). Macrosomia, obesity, macrocephaly and ocular abnormalities (MOMO syndrome) in two unrelated patients: delineation of a newly recognized overgrowth syndrome. *Am. J. Med. Genet.* 46, 555–558. doi: 10.1002/ajmg.1320460519
- Morgan, T. M. (2019). Liquid biopsy: where did it come from, what is it, and where is it going? *Investig. Clin. Urol.* 60, 139–141. doi: 10.4111/icu.2019.60.3.139
- Nakamura, Y., Takahashi, N., Kakegawa, E., Yoshida, K., Ito, Y., Kayano, H., et al. (2008). The GAS5 (growth arrest-specific transcript 5) gene fuses to BCL6 as a result of t(1;3)(q25;q27) in a patient with B-cell lymphoma. *Cancer Genet. Cytogenet.* 182, 144–149. doi: 10.1016/j.cancergencyto.2008.01.013
- Nakashima, E., Tran, J. R., Welting, T. J., Pruijn, G. J., Hirose, Y., Nishimura, G., et al. (2007). Cartilage hair hypoplasia mutations that lead to RMRP promoter inefficiency or RNA transcript instability. *Am. J. Med. Genet.* 143A, 2675–2681. doi: 10.1002/ajmg.a.32053
- Nan, Y., and Zhang, Y. J. (2018). Antisense phosphorodiamidate morpholino oligomers as novel antiviral compounds. *Front. Microbiol.* 9:750. doi: 10.3389/fmicb.2018.00750
- Ng, S. Y., Bogu, G. K., Soh, B. S., and Stanton, L. W. (2013). The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol. Cell* 51, 349–359. doi: 10.1016/j.molcel.2013.07.017
- Ng, M., Heckl, D., and Klusmann, J. H. (2019). The regulatory roles of long noncoding RNAs in acute myeloid leukemia. *Front. Oncol.* 9:570. doi: 10.3389/fonc.2019.00570
- Nguyen, Q., and Carninci, P. (2016). Expression specificity of disease-associated lncRNAs: toward personalized medicine. *Curr. Top. Microbiol. Immunol.* 394, 237–258. doi: 10.1007/82_2015_464
- Ni, N., Song, H., Wang, X., Xu, X., Jiang, Y., and Sun, J. (2017). Up-regulation of long noncoding RNA FALEC predicts poor prognosis and promotes melanoma cell proliferation through epigenetically silencing p21. *Biomed. Pharmacother.* 96, 1371–1379. doi: 10.1016/j.biopha.2017.11.060

- Nik-Zainal, S., Memari, Y., and Davies, H. R. (2020). Holistic cancer genome profiling for every patient. *Swiss Med. Wkly.* 150:w20158. doi: 10.4414/smw.2020.20158
- Ning, S., Yue, M., Wang, P., Liu, Y., Zhi, H., Zhang, Y., et al. (2017). LincSNP 2.0: an updated database for linking disease-associated SNPs to human long non-coding RNAs and their TFBSs. *Nucleic Acids Res.* 45, D74–D78. doi: 10.1093/nar/gkw945
- Pachnis, V., Brannan, C. I., and Tilghman, S. M. (1988). The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO J.* 7, 673–681. doi: 10.1002/j.1460-2075.1988.tb02862.x
- Padua, D., Mahurkar-Joshi, S., Law, I. K., Polytarchou, C., Vu, J. P., Pisegna, J. R., et al. (2016). A long noncoding RNA signature for ulcerative colitis identifies IFNG-AS1 as an enhancer of inflammation. *Am. J. Physiol. Gastrointest. Liver Physiol.* 311, G446–G457. doi: 10.1152/ajpgi.00212.2016
- Panzitt, K., Tschernatsch, M. M., Guelly, C., Moustafa, T., Stradner, M., Strohmaier, H. M., et al. (2007). Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* 132, 330–342. doi: 10.1053/j.gastro.2006.08.026
- Pasic, I., Shlien, A., Durbin, A. D., Stavropoulos, D. J., Baskin, B., Ray, P. N., et al. (2010). Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Res.* 70, 160–171. doi: 10.1158/0008-5472.CAN-09-1902
- Pasmant, E., Laurendeau, I., Heron, D., Vidaud, M., Vidaud, D., and Bieche, I. (2007). Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* 67, 3963–3969. doi: 10.1158/0008-5472.CAN-06-2004
- Pasmant, E., Sabbagh, A., Vidaud, M., and Bieche, I. (2011). ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J.* 25, 444–448. doi: 10.1096/fj.10-172452
- Peng, L., Jiang, J., Tang, B., Nice, E. C., Zhang, Y. Y., and Xie, N. (2020). Managing therapeutic resistance in breast cancer: from the lncRNAs perspective. *Theranostics* 10, 10360–10377. doi: 10.7150/thno.49922
- Petermann, F., Pekowska, A., Johnson, C. A., Jankovic, D., Shih, H. Y., Jiang, K., et al. (2019). The magnitude of IFN- γ responses is fine-tuned by DNA architecture and the non-coding transcript of Ifng-as1. *Mol. Cell* 75, 1229.e1225–1242.e1225. doi: 10.1016/j.molcel.2019.06.025
- Pleasant, E. D., Stephens, P. J., O'Meara, S., McBride, D. J., Meynert, A., Jones, D., et al. (2010). A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 463, 184–190. doi: 10.1038/nature08629
- Pomerantz, M. M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M. P., Doddapaneni, H., et al. (2009). The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat. Genet.* 41, 882–884. doi: 10.1038/ng.403
- Powell, W. T., Coulson, R. L., Crary, F. K., Wong, S. S., Ach, R. A., Tsang, P., et al. (2013). A Prader-Willi locus lncRNA cloud modulates diurnal genes and energy expenditure. *Hum. Mol. Genet.* 22, 4318–4328. doi: 10.1093/hmg/ddt281
- Prensner, J. R., Iyer, M. K., Balbin, O. A., Dhanasekaran, S. M., Cao, Q., Brenner, J. C., et al. (2011). Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol.* 29, 742–749. doi: 10.1038/nbt.1914
- Qu, X., Du, Y., Shu, Y., Gao, M., Sun, F., Luo, S., et al. (2017). MIAT is a pro-fibrotic long non-coding RNA governing cardiac fibrosis in post-infarct myocardium. *Sci. Rep.* 7:42657. doi: 10.1038/srep42657
- Ramsuran, V., Ewy, R., Nguyen, H., and Kulkarni, S. (2018). Variation in the untranslated genome and susceptibility to infections. *Front. Immunol.* 9:2046. doi: 10.3389/fimmu.2018.02046
- Redis, R. S., Vela, L. E., Lu, W., Ferreira de Oliveira, J., Ivan, C., Rodriguez-Aguayo, C., et al. (2016). Allele-specific reprogramming of cancer metabolism by the long non-coding RNA CCAT2. *Mol. Cell* 61, 520–534. doi: 10.1016/j.molcel.2016.01.015
- Ren, C., Liu, F., Ouyang, Z. Y., An, G. L., Zhao, C. H., Shuai, J., et al. (2017). Functional annotation of structural ncRNAs within enhancer RNAs in the human genome: implications for human disease. *Sci. Rep.* 7:15518. doi: 10.1038/s41598-017-15822-7
- Rheinbay, E., Nielsen, M. M., Abascal, F., Wala, J. A., Shapira, O., Tiao, G., et al. (2020). Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* 578, 102–111. doi: 10.1038/s41586-020-1965-x
- Rheinbay, E., Parasuraman, P., Grimsby, J., Tiao, G., Engreitz, J. M., Kim, J., et al. (2017). Recurrent and functional regulatory mutations in breast cancer. *Nature* 547, 55–60. doi: 10.1038/nature22992
- Ridanpaa, M., van Eenennaam, H., Pelin, K., Chadwick, R., Johnson, C., Yuan, B., et al. (2001). Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell* 104, 195–203. doi: 10.1016/S0092-8674(01)00205-7
- Rinn, J. L., and Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166. doi: 10.1146/annurev-biochem-051410-092902
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S. A., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311–1323. doi: 10.1016/j.cell.2007.05.022
- Rochweg, B., Oczkowski, S. J., Siemieniuk, R. A. C., Agoritsas, T., Belley-Cote, E., D'Aragon, F., et al. (2018). Corticosteroids in sepsis: an updated systematic review and meta-analysis. *Crit. Care Med.* 46, 1411–1420. doi: 10.1097/CCM.0000000000003262
- Sacconi, S., Salviati, L., and Desnuelle, C. (2015). Facioscapulohumeral muscular dystrophy. *Biochim. Biophys. Acta* 1852, 607–614. doi: 10.1016/j.bbdis.2014.05.021
- Sahoo, T., del Gaudio, D., German, J. R., Shinawi, M., Peters, S. U., Person, R. E., et al. (2008). Prader-Willi phenotype caused by paternal deficiency for the HBI1-85 C/D box small nucleolar RNA cluster. *Nat. Genet.* 40, 719–721. doi: 10.1038/ng.158
- Samukawa, M., Hirano, M., Saigoh, K., Kawai, S., Hamada, Y., Takahashi, D., et al. (2019). PSP-phenotype in SCA8: case report and systemic review. *Cerebellum* 18, 76–84. doi: 10.1007/s12311-018-0955-0
- Schneider, C., King, R. M., and Philipson, L. (1988). Genes specifically expressed at growth arrest of mammalian-cells. *Cell* 54, 787–793. doi: 10.1016/S0092-8674(88)91065-3
- Schulte, L. N., Schweinlin, M., Westermann, A. J., Janga, H., Santos, S. C., Appenzeller, S., et al. (2020). An advanced human intestinal coculture model reveals compartmentalized host and pathogen strategies during *Salmonella* infection. *mBio* 11:e03348–19. doi: 10.1128/mBio.03348-19
- Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L., et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316, 1341–1345. doi: 10.1126/science.1142382
- Shah, M. Y., Ferracin, M., Pileczki, V., Chen, B., Redis, R., Fabris, L., et al. (2018). Cancer-associated rs6983267 SNP and its accompanying long noncoding RNA CCAT2 induce myeloid malignancies via unique SNP-specific RNA mutations. *Genome Res.* 28, 432–447. doi: 10.1101/gr.225128.117
- Shah, N., and Sukumar, S. (2010). The Hox genes and their roles in oncogenesis. *Nat. Rev. Cancer* 10, 361–371. doi: 10.1038/nrc2826
- Shaker, O. G., Senousy, M. A., and Elbaz, E. M. (2017). Association of rs6983267 at 8q24, HULC rs7763881 polymorphisms and serum lncRNAs CCAT2 and HULC with colorectal cancer in Egyptian patients. *Sci. Rep.* 7:16246. doi: 10.1038/s41598-017-16500-4
- Shen, X., and Corey, D. R. (2018). Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs. *Nucleic Acids Res.* 46, 1584–1600. doi: 10.1093/nar/gkx1239
- Sherr, C. J. (2012). Ink4-Arf locus in cancer and aging. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 731–741. doi: 10.1002/wdev.40
- Siva, N. (2015). UK gears up to decode 100,000 genomes from NHS patients. *Lancet* 385, 103–104. doi: 10.1016/S0140-6736(14)62453-3
- Sivapalaratnam, S., and Bioresource, N. (2018). The rare diseases pilot for the 100,000 genomes project: findings in known and new genes by analysis of 3,549 whole genome sequenced samples from patients and relatives with haematological, haemostasis and immune disorders. *Blood* 132:504. doi: 10.1182/blood-2018-99-119068
- Soibam, B. (2017). Super-lncRNAs: identification of lncRNAs that target super-enhancers via RNA: DNA:DNA triplex formation. *RNA* 23, 1729–1742. doi: 10.1261/rna.061317.117
- Son, H. J., Mo, H. Y., Choi, E. J., Yoo, N. J., and Lee, S. H. (2019). Promoter mutation analysis of long-non-coding RNA RMRP gene in solid tumors. *Pathol. Oncol. Res.* 26, 2809–2810. doi: 10.1007/s12253-018-0530-5
- Sone, M., Hayashi, T., Tarui, H., Agata, K., Takeichi, M., and Nakagawa, S. (2007). The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear

- domain in a subset of neurons. *J. Cell Sci.* 120, 2498–2506. doi: 10.1242/jcs.009357
- Spans, L., Fletcher, C. D., Antonescu, C. R., Rouquette, A., Coindre, J. M., Sciôt, R., et al. (2016). Recurrent MALAT1-GLI1 oncogenic fusion and GLI1 up-regulation define a subset of plexiform fibromyxoma. *J. Pathol.* 239, 335–343. doi: 10.1002/path.4730
- Stacey, S. N., Sulem, P., Masson, G., Gudjonsson, S. A., Thorleifsson, G., Jakobsdottir, M., et al. (2009). New common variants affecting susceptibility to basal cell carcinoma. *Nat. Genet.* 41, 909–914. doi: 10.1038/ng.412
- Stamou, M., Ng, S. Y., Brand, H., Wang, H., Plummer, L., Best, L., et al. (2020). A balanced translocation in Kallmann syndrome implicates a long noncoding RNA, RMST, as a GnRH neuronal regulator. *J. Clin. Endocrinol. Metab.* 105, e231–e244. doi: 10.1210/clinem/dgz011
- Stelzer, Y., Sagi, I., Yanuka, O., Eiges, R., and Benvenisty, N. (2014). The noncoding RNA IPW regulates the imprinted DLK1-DIO3 locus in an induced pluripotent stem cell model of Prader-Willi syndrome. *Nat. Genet.* 46, 551–557. doi: 10.1038/ng.2968
- Stricker, S., and Mundlos, S. (2011). Mechanisms of digit formation: human malformation syndromes tell the story. *Dev. Dyn.* 240, 990–1004. doi: 10.1002/dvdy.22565
- Sung, Y. J., Winkler, T. W., de Las Fuentes, L., Bentley, A. R., Brown, M. R., Kraja, A. T., et al. (2018). A large-scale multi-ancestry genome-wide study accounting for smoking behavior identifies multiple significant loci for blood pressure. *Am. J. Hum. Genet.* 102, 375–400. doi: 10.1016/j.ajhg.2018.01.015
- Tam, V., Patel, N., Turcotte, M., Bosse, Y., Pare, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi: 10.1038/s41576-019-0127-1
- Tang, Q., and Hann, S. S. (2018). HOTAIR: an oncogenic long non-coding RNA in human cancer. *Cell. Physiol. Biochem.* 47, 893–913. doi: 10.1159/000490131
- Tang, J., Li, Y., Sang, Y., Yu, B., Lv, D., Zhang, W., et al. (2018). lncRNA PVT1 regulates triple-negative breast cancer through KLF5/beta-catenin signaling. *Oncogene* 37, 4723–4734. doi: 10.1038/s41388-018-0310-4
- Teppo, S., Laukkanen, S., Liuksiala, T., Nordlund, J., Oittinen, M., Teittinen, K., et al. (2016). Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia. *Genome Res.* 26, 1468–1477. doi: 10.1101/gr.193649.115
- Tragante, V., Barnes, M. R., Ganesh, S. K., Lanktree, M. B., Guo, W., Franceschini, N., et al. (2014). Gene-centric meta-analysis in 87,736 individuals of European ancestry identifies multiple blood-pressure-related loci. *Am. J. Hum. Genet.* 94, 349–360. doi: 10.1016/j.ajhg.2013.12.016
- Tseng, Y. Y., Moriarity, B. S., Gong, W., Akiyama, R., Tiwari, A., Kawakami, H., et al. (2014). PVT1 dependence in cancer with MYC copy-number increase. *Nature* 512, 82–86. doi: 10.1038/nature13311
- Tsujii, H., Yoshimoto, R., Hasegawa, Y., Furuno, M., Yoshida, M., and Nakagawa, S. (2011). Competition between a noncoding exon and introns: gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. *Genes Cells* 16, 479–490. doi: 10.1111/j.1365-2443.2011.01502.x
- Turnbull, C., Ahmed, S., Morrison, J., Pernet, D., Renwick, A., Maranian, M., et al. (2010). Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat. Genet.* 42, 504–507. doi: 10.1038/ng.586
- Tye-Din, J. A., Galipeau, H. J., and Agardh, D. (2018). Celiac disease: a review of current concepts in pathogenesis, prevention, and novel therapies. *Front. Pediatr.* 6:350. doi: 10.3389/fped.2018.00350
- van Dijk, M., Thulluru, H. K., Mulders, J., Michel, O. J., Poutsma, A., Windhorst, S., et al. (2012). HELLP babies link a novel lincRNA to the trophoblast cell cycle. *J. Clin. Invest.* 122, 4003–4011. doi: 10.1172/JCI65171
- van Dijk, M., Visser, A., Buabeng, K. M., Poutsma, A., van der Schors, R. C., and Oudejans, C. B. (2015). Mutations within the LINC-HELLP non-coding RNA differentially bind ribosomal and RNA splicing complexes and negatively affect trophoblast differentiation. *Hum. Mol. Genet.* 24, 5475–5485. doi: 10.1093/hmg/ddv274
- Verhaegh, G. W., Verkleij, L., Vermeulen, S. H., den Heijer, M., Witjes, J. A., and Kiemeny, L. A. (2008). Polymorphisms in the H19 gene and the risk of bladder cancer. *Eur. Urol.* 54, 1118–1126. doi: 10.1016/j.eururo.2008.01.060
- Vigneau, S., Rohrlrich, P. S., Brahic, M., and Bureau, J. F. (2003). Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *J. Virol.* 77, 5632–5638. doi: 10.1128/JVI.77.10.5632-5638.2003
- Volders, P. J., Lefever, S., Baute, S., Nuytens, J., Vanderheyden, K., Menten, B., et al. (2018). Targeted genomic screen reveals focal long non-coding RNA copy number alterations in cancer cell lines. *Noncoding RNA* 4:21. doi: 10.3390/ncrna4030021
- Vu, P. Y., Toutain, J., Cappellen, D., Delrue, M. A., Daoud, H., El Moneim, A. A., et al. (2012). A homozygous balanced reciprocal translocation suggests LINC00237 as a candidate gene for MOMO (macrosomia, obesity, macrocephaly, and ocular abnormalities) syndrome. *Am. J. Med. Genet.* 158A, 2849–2856. doi: 10.1002/ajmg.a.35694
- Wahba, A. S., Ibrahim, M. E., Mesbah, N. M., Saleh, S. M., Abo-Elmatty, D. M., and Mehanna, E. T. (2020). Serum LINC00305 expression and its genetic variant rs2850711 are associated with clinical and laboratory features of rheumatoid arthritis. *Br. J. Biomed. Sci.* 77, 142–147. doi: 10.1080/09674845.2020.1744942
- Walavalkar, K., Saravanan, B., Singh, A. K., Jayani, R. S., Nair, A., Farooq, U., et al. (2020). A rare variant of African ancestry activates 8q24 lncRNA hub by modulating cancer associated enhancer. *Nat. Commun.* 11:3598. doi: 10.1038/s41467-020-17325-y
- Wang, Y., Guo, Z., Zhao, Y., Jin, Y., An, L., Wu, B., et al. (2017). Genetic polymorphisms of lncRNA-p53 regulatory network genes are associated with concurrent chemoradiotherapy toxicities and efficacy in nasopharyngeal carcinoma patients. *Sci. Rep.* 7:8320. doi: 10.1038/s41598-017-08890-2
- Wang, H., Li, W., Guo, R., Sun, J., Cui, J., Wang, G., et al. (2014). An intragenic long noncoding RNA interacts epigenetically with the RUNX1 promoter and enhancer chromatin DNA in hematopoietic malignancies. *Int. J. Cancer* 135, 2783–2794. doi: 10.1002/ijc.28922
- Wevrick, R., Kerns, J. A., and Francke, U. (1994). Identification of a novel paternally expressed gene in the Prader-Willi syndrome region. *Hum. Mol. Genet.* 3, 1877–1882. doi: 10.1093/hmg/3.10.1877
- Xiang, J. F., Yin, Q. F., Chen, T., Zhang, Y., Zhang, X. O., Wu, Z., et al. (2014). Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* 24, 513–531. doi: 10.1038/cr.2014.35
- Xin, Y., Li, Z., Zheng, H., Chan, M. T. V., and Wu, W. K. K. (2017). CCAT2: a novel oncogenic long non-coding RNA in human cancers. *Cell Prolif.* 50:e12342. doi: 10.1111/cpr.12342
- Xu, Y., and Shao, B. (2018). Circulating lncRNA IFNG-AS1 expression correlates with increased disease risk, higher disease severity and elevated inflammation in patients with coronary artery disease. *J. Clin. Lab. Anal.* 32:e22452. doi: 10.1002/jcla.22452
- Xu, T., Yan, S., Jiang, L., Yu, S., Lei, T., Yang, D., et al. (2019). Gene amplification-driven long noncoding RNA SNHG17 regulates cell proliferation and migration in human non-small-cell lung cancer. *Mol. Ther. Nucleic Acids* 17, 405–413. doi: 10.1016/j.omtn.2019.06.008
- Yamagata, K., Nakayamada, S., and Tanaka, Y. (2020). Critical roles of super-enhancers in the pathogenesis of autoimmune diseases. *Inflamm. Regen.* 40:16. doi: 10.1186/s41232-020-00124-9
- Yang, C., Tang, R., Ma, X., Wang, Y., Luo, D., Xu, Z., et al. (2015). Tag SNPs in long non-coding RNA H19 contribute to susceptibility to gastric cancer in the Chinese Han population. *Oncotarget* 6, 15311–15320. doi: 10.18632/oncotarget.3840
- Yang, Z., Zhao, S., Zhou, X., Zhao, H., and Jiang, X. (2019). PCAT-1: a pivotal oncogenic long non-coding RNA in human cancers. *Biomed. Pharmacother.* 110, 493–499. doi: 10.1016/j.biopha.2018.12.014
- Yap, K. L., Li, S., Munoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., et al. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* 38, 662–674. doi: 10.1016/j.molcel.2010.03.021
- Yu, X., Zheng, H., Chan, M. T., and Wu, W. K. (2017). HULC: an oncogenic long non-coding RNA in human cancer. *J. Cell. Mol. Med.* 21, 410–417. doi: 10.1111/jcmm.12956
- Yuan, L., Xu, Z. Y., Ruan, S. M., Mo, S., Qin, J. J., and Cheng, X. D. (2020). Long non-coding RNAs towards precision medicine in gastric cancer: early diagnosis, treatment, and drug resistance. *Mol. Cancer* 19:96. doi: 10.1186/s12943-020-01219-0
- Zack, T. I., Schumacher, S. E., Carter, S. L., Cherniack, A. D., Saksena, G., Tabak, B., et al. (2013). Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* 45, 1134–1140. doi: 10.1038/ng.2760
- Zeggini, E., Weedon, M. N., Lindgren, C. M., Frayling, T. M., Elliott, K. S., Lango, H., et al. (2007). Replication of genome-wide association signals in

- UK samples reveals risk loci for type 2 diabetes. *Science* 316, 1336–1341. doi: 10.1126/science.1142364
- Zhang, X., Hamblin, M. H., and Yin, K. J. (2017c). The long noncoding RNA Malat1: its physiological and pathophysiological functions. *RNA Biol.* 14, 1705–1714. doi: 10.1080/15476286.2017.1358347
- Zhang, B. Y., Jin, Z., and Zhao, Z. (2017a). Long intergenic noncoding RNA 00305 sponges miR-136 to regulate the hypoxia induced apoptosis of vascular endothelial cells. *Biomed. Pharmacother.* 94, 238–243. doi: 10.1016/j.biopha.2017.07.099
- Zhang, D. D., Wang, W. T., Xiong, J., Xie, X. M., Cui, S. S., Zhao, Z. G., et al. (2017b). Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NF- κ B pathway in human monocytes. *Sci. Rep.* 7:46204. doi: 10.1038/srep46204
- Zhang, X., Zhou, L., Fu, G., Sun, F., Shi, J., Wei, J., et al. (2014). The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. *Carcinogenesis* 35, 2062–2067. doi: 10.1093/carcin/bgu103
- Zhao, J., Du, P., Cui, P., Qin, Y., Hu, C., Wu, J., et al. (2018). LncRNA PVT1 promotes angiogenesis via activating the STAT₃/VEGFA axis in gastric cancer. *Oncogene* 37, 4094–4109. doi: 10.1038/s41388-018-0250-z
- Zhao, R., Sun, F., Bei, X., Wang, X., Zhu, Y., Jiang, C., et al. (2017). Upregulation of the long non-coding RNA FALEC promotes proliferation and migration of prostate cancer cell lines and predicts prognosis of PCa patients. *Prostate* 77, 1107–1117. doi: 10.1002/pros.23367
- Zhou, X., Han, X., Wittfeldt, A., Sun, J., Liu, C., Wang, X., et al. (2016b). Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF- κ B pathway. *RNA Biol.* 13, 98–108. doi: 10.1080/15476286.2015.1122164
- Zhou, C. C., Yang, F., Yuan, S. X., Ma, J. Z., Liu, F., Yuan, J. H., et al. (2016a). Systemic genome screening identifies the outcome associated focal loss of long noncoding RNA PRAL in hepatocellular carcinoma. *Hepatology* 63, 850–863. doi: 10.1002/hep.28393
- Zhou, W., Zhao, Z., Wang, R., Han, Y., Wang, C., Yang, F., et al. (2017). Identification of driver copy number alterations in diverse cancer types and application in drug repositioning. *Mol. Oncol.* 11, 1459–1474. doi: 10.1002/1878-0261.12112
- Zhou, B., Zhao, H., Yu, J., Guo, C., Dou, X., Song, F., et al. (2018). EVLncRNAs: a manually curated database for long non-coding RNAs validated by low-throughput experiments. *Nucleic Acids Res.* 46, D100–D105. doi: 10.1093/nar/gkx677
- Zhu, M., Li, N., Luo, P., Jing, W., Wen, X., Liang, C., et al. (2018). Peripheral blood leukocyte expression of lncRNA miat and its diagnostic and prognostic value in ischemic stroke. *J. Stroke Cerebrovasc. Dis.* 27, 326–337. doi: 10.1016/j.jstrokecerebrovasdis.2017.09.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Aznaourova, Schmerer, Schmeck and Schulte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.