Exploring the world of non-coding genes in stem cells and autoimmunity.

Messemaker, T.C.

Chapter 8
Summarizing discussion

Numerous studies have contributed to our current understanding of autoimmune diseases (AIDs), however, pathogenesis of many AIDs can still not be fully explained. Both genetic factors and environmental factors are involved in the onset of autoimmunity. Which mechanisms explain the contribution of these genetic and environmental factors to disease pathogenesis, and how the different factors interplay remain unanswered key questions. The studies presented in this thesis aimed at identifying and unravelling some of the enigmatic mechanisms in rheumatoid arthritis (RA) and systemic sclerosis (SSc).

Epigenetic changes are thought to play a role in passing on environmental influences to gene expression alterations that can contribute to disease. In **Chapter 2** of this thesis, we investigated whether monocytes from diagnosed but yet untreated RA patients contain a distinct, disease-related, epigenetic signature of genes associated with RA. No large epigenetic differences were observed between RA and healthy monocytes indicating that such differences are either small or not present in the tested cell type for TNFα and IL6. However, epigenetic differences were observed in RA patients in other cell types indicating that epigenetic changes can play an important role in RA. DNA hypomethylation was found in synovial fibroblast from RA patients and indicate that cells in a disease-affected environment may display epigenetic differences[1]. Therefore, also monocytes or other cell types in the synovium may display differences and thereby contribute to disease pathogenesis. It would be interesting to investigate whether epigenetic differences are also present and maintained in precursor cells (like CD34+ cells). Upon differentiation, these cells may end up in the joints and trigger or enhance the inflammatory status found in RA patients[2,3]. Which cell types do contain these epigenetic traces, how they obtain these marks and how we can restore an autoimmune epigenetic landscape are topics for future studies. Moreover, another key question is whether these epigenetic marks are present prior to the onset of the disease or whether these marks are a consequence of disease pathogenesis. This question may be answered by longitudinal retrospective studies in which the responsible cell types have been collected. In case these marks are present prior to the onset of a disease, they may also play a crucial role in the diagnoses and treatment of autoimmunity by opening early treatment options. Together, further efforts investigating how and which epigenetic changes are involved in disease pathogenesis on a genome-wide level and a cell-type specific manner are needed to increase our

understanding in disease pathogenesis and may reveal early diagnostic markers or open up novel treatment options.

Large genetic studies containing the genetic information of over 100.000 individuals have been performed to relate variants and genes to a role in disease pathogenesis of rheumatoid arthritis[4]. These genetic population studies can identify hundreds of variants in a single locus that all associate with disease due to high linkage disequilibrium. Identifying the causal SNPs is often difficult as the highest associated variant (lead SNP) of a disease associated locus is not necessarily the causing variant[5]. Revealing which functional mechanisms shelter behind associated SNPs aids in understanding how genes are affected and which pathways may play a role in disease pathogenesis. **Chapter 3** of this thesis reviews identified variants contributing functionally to disease, and the involved pathways that are hypothesized to play a role in RA. For example, the coding variant (Arg620Trp) in PTPN22 was shown to affect both BCR and TCR signalling[6]. Moreover, several variants in different genes have shown to affect NF-kB signalling, including: the variants Val194Ala and Pro175Leu in NFKBIE, variant Phe127Cys in TNFA3 and variant Ala288Thr in RTKN2. Similar evidence for the involvement of these pathways came forward from gene enrichment analysis of candidate genes located in the 100 associated risk loci which identified T-cell receptor (TCR) signalling, NF-κb signalling and JAK-STAT signalling as the most enriched processes (**Chapter 3**)[6]. Several other studies have investigated the role of these pathways in context of autoimmunity[7,8]. In the JAK-STAT signalling cascade, STAT is phosphorylated by JAK proteins resulting in the activation of pro-inflammatory cytokines thereby promoting the inflammatory state in RA patients[9]. Inhibitors of this cascade have with success been tested as therapeutics reducing the level of pro inflammatory cytokines[10]. Tofacitinib, a JAK-STAT inhibitor has received FDA approval and several other inhibitors are being tested in clinical trials[11–13]. Similarly, functional studies have highlighted enhanced NF-κB activity and defective TCR signalling in RA patients[1415]. Studies are undergoing investigating potential therapeutics targeting both TCR receptor signalling and the NF-κB cascade[16–19]. Together, we hypothesize that non-HLA RA-associated variants in these genes and pathway are responsible for a decreased immune activation threshold and for disturbing a healthy ratio between pro and anti-inflammatory cytokines increasing the probability of developing RA. Although for some RA-associated variants the casual mechanisms has been revealed, future

studies should be conducted for the remaining variants. Thereby, understanding the influence of variants in these genes from the identified pathways might also explain why some of the used therapeutics is not beneficial for all RA patients and stimulates the research into personal medicine within the field of RA and autoimmunity.

However for the majority of risk loci, the causal mechanisms for their association with RA remain elucidative. One of these loci is the *TRAF1-C5* locus which contains multiple RA-associating variants in high linkage disequilibrium of which the causal variant has not yet been identified. Although variants in *C5* have been identified as variants affecting C5 function, these variants do not significantly associate with RA. It is therefore unlikely that these variants can explain the association of this region with RA as described in **Chapter 4**. The *TRAF1-C5* locus lacks RA-associated variants that change amino acids of the nearby candidate genes and therefore no functional mechanism have been identified. In **Chapter 5** of this thesis we describe our discovery of a novel gene named *C5T1lncRNA* in this region. Interestingly, two SNPs are located in the RNA sequence of this presumably long non-coding RNA (lncRNA). Non-coding RNAs do not translate into proteins and these SNPs are therefore not identified as amino acid changing variants. Nonetheless, SNPs in lncRNAs can be functional variants as several studies have shown that SNPs can alter i.) the binding potential of the lncRNA, ii.) the structure of the lncRNA and iii.) lncRNA expression levels[20–22]. Moreover, the identified lncRNA in the *TRAF1-C5* locus was found to be expressed and functional in RA-relevant cell types as synovial fibroblasts and may therefore have a functional role in RA pathogenesis. We speculate a mechanism in which variants in *C5T1-lncRNA* might interfere with the function of this gene. In **Chapter 5**, we found that decreasing levels of C5T1lncRNA also decreased levels of the nearby gene *C5* indicating a regulatory role. Variants in *C5T1lncRNA* might therefore interfere with this regulatory role and might thus also affect the function of C5, a potent pro-inflammatory immune gene. Future studies should be designed to investigate the effect of the variants in the *TRAF1-C5* locus and what consequences this brings for *C5*. Thereby, we cannot rule out the possibility that variants in the *TRAF1-C5* region influences either with *C5* and *TRAF1* via other mechanisms. Several eQTL effects were found from variants in the *TRAF1-C5* region[23,24]. These variants could interfere with C5 and TRAF1 levels by for example influencing the mRNA stability or by interfering with transcription

factors binding sites. Such mechanisms could function as causal mechanisms for RA independent of *C5T1lncRNA*. Additionally, a cell-type specific manner in which variants affect genes in the *TRAF1-C5* locus is possible[25]. *C5T1lncRNA* is highly expressed in the liver, similar to *C5*, but *C5T1lncRNA* is also strongly induced by LPS in monocytes, similar to *TRAF1*, illustrating the complex nature of this locus[26]. In order to aid in addressing the functional mechanisms of such loci, large studies have been set up to collect cell type specific expression in hundreds of cell types[27]. Currently, FANTOM5, TiGER and GTEX are large databases that provide such expression data of over 20.000 genes in more than 400 cell types and over 100 different tissues providing useful platforms for future expression studies[27–29].

To identify functional variants originating from genome wide association studies (GWAS) and to understand genomic variation, large studies have been set up focussing on gene expression changes linked to genomic variation, also known as eQTL studies. A large study that included over 5000 individuals identified that genetic variations can influence gene expression of genes, both in *cis* and in *trans*[30]. Another large study investigated expression changes specifically in monocytes from over 1000 individuals and reported similar findings[31]. These studies provide a useful platform and starting point for the unravelling of functional genetic variants. Such studies also provide insight into which cell types play a role in disease by investigating cell-type specific eQTLs. A recent study investigated cell type specific eQTLs in monocytes and B-cells and showed that disease associating variants can have functional consequences in a cell type specific manner[32]. Moreover, Raj *et al*. investigated cell type specific traits in T-cells and monocytes and identified that many variants associated with RA specifically influenced the expression of genes in T-cells[33]. From these studies it has been concluded that variants often display cell specific traits and may indicate which cell types play a role in disease pathogenesis. Additional genetic evidence showed that T-cells play an important role in RA. Overlapping disease-associating variants with the presence of active or repressing histone modifications in a cell type specific manner provides indications in which cell type, which variants are being accessible. Farh *et al*. found that RA-associating variants display histone modifications that are enriched in T-cells, B-cells and lymphoblastoid cells in a comparison with 33 different cell-types[34]. Finally, examining lncRNA expression in RA-associated loci has been linked to T-cells as Hrdlicknova *et al*. has shown that the lncRNAs located in associated regions are

often specifically expressed in T-cells[35]. These studies illustrate which cell types may be responsible and indicate that not only coding genes but also non-coding genes are potential disease genes when affected by variants. Although enrichment statistics and gene coexpression are not conclusive with regard to causality and functionality, additional functional studies are necessary. Nonetheless, it is likely that development to RA is affected by defects in multiple cell types of which T cells and T-cell activation play an important and determining role. Genetic variants likely affect genes in a cell specific manner resulting together with other cellular defects and environmental alterations in an increased susceptibility to RA.

Aside from genetic studies, RNA sequencing of disease-relevant tissues can also highlight genes and pathways involved in disease pathogenesis. In **Chapter 6,** the RNA of skin from SSc patients was compared with skin from healthy donors, and resulted in the identification of both deregulated coding and non-coding genes. In this chapter specifically non-coding genes were investigated and hundreds of deregulated lncRNAs were observed. Among these, several lncRNAs were validated using a replication dataset, including AGAP2-AS1, CTBP1-AS2 and OTUD6B-AS1. These genes are classified as antisense genes and in both studies, also their sense gene was deregulated. Although no functional assessment was performed in this study, we hypothesize that such deregulated gene pairs play a role in the disease pathogenesis of SSc. In such a model the deregulated antisense gene fails to maintain its regulatory role on its opposing sense gene resulting in a deregulated gene pair leading to depending on its function to disease pathogenesis. Coinciding with this model is the high correlation that was found between the expression of both genes within such gene pairs in our study and other studies[36–38]. Overall, we hypothesize that some of these lncRNAs either are involved with functions contributing to SSc directly, or by influencing other coding genes thereby contributing to SSc pathogenesis. Although thousands of long non-coding RNAs have been discovered, very few molecular mechanisms have yet been identified. lncRNAs can have a diverse set of functions and interfere not only in disease pathogenesis but also developmental processes. Like described in **Chapter 7**, *Sox2ot*, a lncRNA overlapping *Sox2*, interferes with *Sox2* gene transcription. Sox2ot is a gene that is located near enhancer and transcription regions that are important for Sox2 expression. Expression of Sox2ot is hypothesized to interfere with the transcriptional process of Sox2

thereby regulating its levels. In a developmental point of view, similar mechanisms are possible for other development genes. For example, Sox1 and Sox4 display a similar genetic landscape and might therefore also be under regulation of non-coding RNA transcripts. The hypothesized mechanism of Sox2ot that came forward from the study in **Chapter 7** was interference of enhancer regions by altering DNA-looping events. Currently, studies are on-going to reveal in-depth genetic landscapes and cross-communication of genes, enhancers, transcription factors, via chromatin-loops[39,40]. Novel methods allow more detailed overview of this genetic landscape and will aid in unravelling non-coding RNA functions and disease mechanisms. Together, our studies contribute to a better understanding of how genes are regulated, which DNA regions are responsible for gene activation and gene silencing and whether non-coding genes might be involved.

Unravelling the function of lncRNAs is essential to understand their role and involvement in development but also in diseases like autoimmunity. Currently several laboratories have set up large scale experiments to investigate these functions, especially in cancer by evaluating lncRNAs involved in cell growth[41,42]. These studies have identified numerous lncRNAs functionally involved in cell growth in several cancer cell lines. However not all lncRNAs function through interference with cell growth and therefore similar studies should be set up focusing on other cellular functions. An example would be to knock down levels of (or knockout) lncRNAs in immune cell types followed by various immune activation signals to identify which lncRNAs are involved in the immune response. In the near future, such studies will be performed and will be aided by the revolution of CRISPR technology allowing largescale knockdown technology.

More and more lncRNAs are being identified as deregulated genes in disease and development which opens the possibility to use them as diagnostic markers or therapeutic targets. Although, non-coding genes are overall lower expressed compared to coding genes, they also possess characteristics that will prefer non-coding genes over coding genes as future drug targets. For example their cell-type specificity allows drugs to be effective in one cell-type only, preventing unwanted side effects in other cell types or tissues. Especially in cancers, where cancer-specific lncRNA expression can be used as a therapeutic targets thereby leaving healthy tissue unaffected. The first report has already shown that targeting a lncRNA known as MALAT by antisense oligo nucleotides was able to

prevent lung cancer metastasis in mice displaying the feasibility of targeting lncRNAs[43]. Other potential intervention approaches through lncRNAs that are in pre-clinical development include siRNAs, aptamers, ribozymes or small molecules and are reviewed in ref[44]. As lncRNAs are often highly expressed in specific diseased cells (like cancer cells) they can also be used as biomarkers and for diagnostic purposes. A diagnostic test using an overexpressed lncRNA is currently under development and is applicable for the diagnoses of prostate cancer[45]. This test can measure levels of PCA3, a prostate specific lncRNA overexpressed in prostate cancer, in the urine of patients[45]. With rapidly advancing technology it will be easier to detect and target lncRNAs and therefore an increasing amount of specific biomarkers for early diagnoses, better prognostic prediction and more efficient therapy will undoubtedly be available in future clinical applications.

The studies presented in this thesis contributed to the identification of lncRNAs involved in disease pathogenesis. Although non-coding RNAs are overall lower expressed, still they may regulate crucial functions and should not be disregarded merely based on present abundances. Future single-cell sequencing studies will be able to gather detailed information regarding non-coding RNAs and their mechanisms in cell specific manners. Together the reducing costs for sequencing, the increasing single cell resolution to study gene expression and the efficient single cell isolation technology provide a highly accurate platform to study both basic and translational research. Expression profiles of both coding and non-coding RNAs on single cell levels may aid in the identification and characterisation of novel and existing cell types. Therefore further unravelling mechanisms by which non-coding RNAs function not only lead to insight in disease development but we hypothesise the idea that non-coding genes will one day be used as target genes in future therapies, including diseases of autoimmunological nature. Finally, if epigenetic alterations (such as histone modifications or non-coding RNA dysregulation) occur years before the onset of a disease, they may be better therapeutic targets prevent the disease compared to current medicines who are often used to supress the disease or to treat the symptoms only.

References

1 Karouzakis E, Gay RE, Michel B a, Gay S, Neidhart M. DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum* 2009; **60**: 3613–22.

2 Hirohata S, Yanagida T, Nampei A, Kunugiza Y, Hashimoto H, Tomita T *et al.* Enhanced generation of endothelial cells from CD34+ cells of the bone marrow in rheumatoid arthritis: Possible role in synovial neovascularization. *Arthritis Rheum* 2004; **50**: 3888–3896.

3 Hirohata S, Yanagida T, Tomita T, Ochi T, Nakamura H, Yoshino S. Bone marrow CD34+ progenitor cells from rheumatoid arthritis patients give rise to spontaneous transformation of peripheral blood B cells from Epstein-Barr virus infected healthy individuals. *Arthritis Rheum* 1996; **39**: 1002.

4 Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2013. doi:10.1038/nature12873.

5 Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome TL - 22. *Genome Res* 2012; **22 VN**-**r**: 1748–1759.

6 Messemaker TC, Huizinga TW, Kurreeman F. Immunogenetics of rheumatoid arthritis: Understanding functional implications. *J Autoimmun* 2015; **64**: 7–14.

7 Banerjee S, Biehl A, Gadina M, Hasni S, Schwartz DM. JAK–STAT Signaling as a Target for Inflammatory and Autoimmune Diseases: Current and Future Prospects. *Drugs* 2017; **77**: 521–546.

8 Makarov SS. NF-kappa B in rheumatoid arthritis: a pivotal regulator of inflammation, hyperplasia, and tissue destruction. *Arthritis Res* 2001; **3**: 200–6.

9 Darnell J, Kerr I, Stark G. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science (80-)* 1994; **264**: 1415–1421.

10 Malemud CJ. Inhibitors of JAK for the treatment of rheumatoid arthritis: rationale and clinical data. *Clin Investig (Lond)* 2012; **2**: 39–47.

11 Kremer JM, Emery P, Camp HS, Friedman A, Wang L, Othman AA *et al.* A Phase IIb Study of ABT-494, a Selective JAK-1 Inhibitor, in Patients With Rheumatoid Arthritis and an Inadequate Response to Anti–Tumor Necrosis Factor Therapy. Arthritis Rheumatol. 2016; **68**: 2867–2877.

12 Boyle DL, Soma K, Hodge J, Kavanaugh A, Mandel D, Mease P *et al.* The JAK inhibitor tofacitinib suppresses synovial JAK1-STAT signalling in rheumatoid arthritis. *Ann Rheum Dis* 2015; **74**: 1311–1316.

13 Vanhoutte F, Mazur M, Voloshyn O, Stanislavchuk M, Van der Aa A, Namour F *et al.* Efficacy, safety, pharmacokinetics, and pharmacodynamics of filgotinib, a selective Janus kinase 1 inhibitor, after short-term treatment of rheumatoid arthritis: Results of two randomized Phase IIA trials. *Arthritis Rheumatol* 2017; **69**: 1949–1959.

14 Maurice MM, Lankester a C, Bezemer a C, Geertsma MF, Tak PP, Breedveld FC *et al.* Defective TCR-mediated signaling in synovial T cells in rheumatoid arthritis. *J Immunol* 1997; **159**: 2973–8.

15 Lawrence T, Gilroy DW, Colville-Nash PR, Willoughby DA. Possible new role for NF-kappaB in the resolution of inflammation. Nat. Med. 2001; **7**: 1291–1297.

16 Esensten JH, Helou YA, Chopra G, Weiss A, Bluestone JA. CD28 Costimulation: From Mechanism to Therapy. *Immunity* 2016; **44**: 973–988.

17 Thomas R, Turner M, Cope AP. High avidity autoreactive T cells with a low signalling capacity through the T-cell receptor: central to rheumatoid arthritis pathogenesis? *Arthritis Res Ther* 2008; **10**: 210.

18 Zhuang Y, Liu J, Ma P, Bai J, Ding Y, Yang H *et al.* Tamarixinin a alleviates joint destruction of rheumatoid arthritis by blockade of MAPK and NF-κB activation. Front. Pharmacol. 2017; **8**. doi:10.3389/fphar.2017.00538.

19 Okamoto H, Yoshio T, Kaneko H, Yamanaka H. Inhibition of NF-KappaB signaling by fasudil as a potential therapeutic strategy for rheumatoid arthritis. *Arthritis Rheum* 2010; **62**: 82–92.

20 Mirza AH, Kaur S, Brorsson CA, Pociot F. Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. *PLoS One* 2014; **9**. doi:10.1371/journal.pone.0105723.

21 Almlöf JC, Lundmark P, Lundmark A, Ge B, Pastinen T, Goodall AH *et al.* Single nucleotide polymorphisms with cis-regulatory effects on long non-coding transcripts in human primary monocytes. *PLoS One* 2014; **9**: e102612.

22 Bhartiya D, Scaria V. Genomics Genomic variations in non-coding RNAs : Structure , function and regulation. *Genomics* 2016; **107**: 59–68.

23 van Steenbergen HW, Rodríguez-Rodríguez L, Berglin E, Zhernakova A, Knevel R, Ivorra-Cortés J *et al.* A genetic study on C5-TRAF1 and progression of joint damage in rheumatoid arthritis. *Arthritis Res Ther* 2015; **17**: 1.

24 Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 2016; **48**: 481–487.

25 Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* 2012; **44**: 502–510.

26 Messemaker TC, Frank-Bertoncelj M, Marques RB, Adriaans a, Bakker a M, Daha N *et al.* A novel long non-coding RNA in the rheumatoid arthritis risk locus TRAF1-C5 influences C5 mRNA levels. *Genes Immun* 2016; **17**: 85–92.

27 Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S *et al.* Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol* 2015; **16**: 22.

28 Liu X, Yu X, Zack DJ, Zhu H, Qian J. TiGER: a database for tissue-specific gene expression and regulation. *BMC Bioinformatics* 2008; **9**: 271.

29 Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**: 580–585.

30 Westra H, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; **45**: 1238–1243.

31 Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R *et al.* Genetics and beyond - the transcriptome of human monocytes and disease susceptibility. *PLoS One* 2010; **5**. doi:10.1371/journal.pone.0010693.

32 Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A *et al.* Genetics of gene expression in primary immune cells identifies cell type – specific master regulators and roles of HLA alleles. *Nat Genet* 2012; **44**: 502–510.

33 Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM *et al.* Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science (80-)* 2014; **344**: 519–23.

34 Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015; **518**: 337–43.

35 Hrdlickova B, Kumar V, Kanduri K, Zhernakova D V, Tripathi S, Karjalainen J *et al.* Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. *Genome Med* 2014; **6**: 88.

36 Goyal A, Fiškin E, Gutschner T, Polycarpou-Schwarz M, Groß M, Neugebauer J *et al.* A cautionary tale of sense-antisense gene pairs: independent regulation despite inverse correlation of expression. *Nucleic Acids Res* 2017; **45**: 12496–12508.

37 Messemaker TC, Chadli L, Cai G, Goelela VS, Boonstra M, Dorjée AL *et al.* Antisense long non-coding RNAs are deregulated in skin tissue of patients with systemic sclerosis. *J Invest Dermatol* 2017. doi:10.1016/j.jid.2017.09.053.

38 Balbin OA, Malik R, Dhanasekaran SM, Prensner JR, Cao X, Wu Y *et al.* The landscape of antisense gene expression in human cancers. *Genome Res* 2015; **25**: 1068–1079.

39 Zhang H, Li F, Jia Y, Xu B, Zhang Y, Li X *et al.* Characteristic arrangement of nucleosomes is predictive of chromatin interactions at kilobase resolution. *Nucleic Acids Res* 2017; : 1–13.

40 Barutcu AR, Fritz AJ, Zaidi SK, van Wijnen AJ, Lian JB, Stein JL, Nickerson JA, Imbalzano AN SG. C-ing the Genome: A Compendium of Chromosome Conformation Capture Methods to Study Higher-Order Chromatin Organization. *J Cell Physiol* 2016; **21**: 133–136.

41 Zhu S, Li W, Liu J, Chen C, Liao Q, Xu P *et al.* Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR–Cas9 library. *Nat Biotechnol* 2016; **34**: 1279–1286.

42 Liu SJ, Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science (80-)* 2016; **7111**: aah7111.

43 Gutschner T, Hämmerle M, Eißmann M, Hsu J, Kim Y, Hung G *et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 2013; **73**: 1180–1189.

44 Parasramka MA, Maji S, Matsuda A, Yan IK, Patel T. Pharmacology & Therapeutics Long non-coding RNAs as novel targets for therapy in hepatocellular carcinoma. *Pharmacol Ther* 2016; **161**: 67–78.

45 Lee GL, Dobi A, Srivastava S. Diagnostic performance of the PCA3 urine test. *Prostate Cancer* 2011; **8**: 123–124.