Trends in Cell Biology

LncRNAs in human cancers: signal from noise

--Manuscript Draft--

Manuscript Number:	TCB-D-21-00196R1
Article Type:	Opinion
Keywords:	LncRNAs; biomolecular condensates; stochasticity; LLPS; membraneless bodies; molecular crowding
Corresponding Author:	Eleonora Leucci Leuven, BELGIUM
First Author:	Eleonora Leucci
Order of Authors:	Eleonora Leucci
	Sara Adnane
	Alessandro Marino
Abstract:	Given the biochemical reactions stochasticity, the mechanisms leading to conservation of biological functions from noise are obscure. Pervasive transcription of non-conserved genomic regions, generates lowly-expressed cancer-specific lncRNAs. How such poorly expressed transcripts, often undetectable in normal tissues, consistently modulate the activity of multiple abundant proteins leading to cancer phenotypes is unclear. Biochemical reaction compartmentalization in response to environmental oscillations through liquid-liquid phase separation (LLPS) may explain the emergence of order from molecular noise. LncRNAs contain repetitive sequences and as such contribute to molecular crowding and LLPS. We propose that lncRNAs mediate cancer stress signals by regulating aberrant LLPS. This emerging model and its consequences for stoichiometry and specificity may lead to the development of diagnostic tools and cancer-specific drugs.

Highlights

1. The human (cancer) genome is pervasively transcribed into a plethora of non-coding transcripts that are mostly not-conserved, lowly expressed and consistently derepressed in cancer.
2. LncRNAs exert multiple key molecular functions in cancer, converging towards the regulation of epigenetic and post- transcriptional events.
3. The lack of conservation and the low expression of lncRNAs are in striking contrast with their key role in all epigenetic and post- transcriptional processes in the cell thus exposing the CC (Conservation and Concentration) paradox.
4. LncRNA are able to drive aberrant LLPS in cancer in response to stressors.
5. A better understanding of lncRNA driven-cancer aberrant compartmentalisation may lead to the development of new diagnostic and therapeutic tools and to the re-evaluation of basic concepts in RNA-protein interactions.

LncRNAs in human cancers: signal from noise

Sara Adnane[1], Alessandro Marino[1] and Eleonora Leucci[1,2,*]

[1]*Laboratory for RNA Cancer Biology, Department of Oncology, KULeuven, Herestraat 49, 3000 Leuven, Belgium.*

[2]*Trace, LKI Leuven Cancer Institute, KU Leuven, Herestraat 49, 3000 Leuven, Belgium*.

Keywords: lncRNAs; biomolecular condensates; stochasticity; LLPS; membraneless bodies; molecular crowding.

*Corresponding Author

Prof. Eleonora Leucci,
Laboratory for RNA Cancer Biology
Department of Oncology, KU Leuven
Campus Gasthuisberg, Onderwijs en Navorsing I
Herestraat 49, bus 818
3000 LEUVEN
eleonora.leucci@kuleuven.be
Phone: +32 16 32 32 16
https://gbiomed.kuleuven.be/english/research/50488876/54502087
Twitter: @LeucciLRCB

Abstract

Given the biochemical reactions stochasticity, the mechanisms leading to conservation of biological functions from noise are obscure. Pervasive transcription of non-conserved genomic regions, generates lowly-expressed cancer-specific lncRNAs. How such poorly expressed transcripts, often undetectable in normal tissues, consistently modulate the activity of multiple abundant proteins leading to cancer phenotypes is unclear. Biochemical reaction compartmentalization in response to environmental oscillations through liquid-liquid phase separation (LLPS) may explain the emergence of order from molecular noise. LncRNAs contain repetitive sequences and as such contribute to molecular crowding and LLPS. We propose that lncRNAs mediate cancer stress signals by regulating aberrant LLPS. This emerging model and its consequences for stoichiometry and specificity may lead to the development of diagnostic tools and cancer-specific drugs.

A role for stochasticity and lncRNAs in the organisation of cancer cells

The increasing use of single cell and single molecule techniques has revealed that all biochemical reactions in cells are intrinsically **stochastic** (see Glossary) [1, 2] . Even a process that has thus far been considered deterministic and unidirectional such as gene expression is permeated by low affinity interactions and stochasticity at all stages from transcription (**pervasive transcription**) to translation (pervasive translation).

Within this overt randomness however, structural organization arises to preserve key biological functions. A recent study has shown that in oscillating systems, random heterogeneity consistently promotes organization and outperforms design in network organization [3]. Accordingly, stochasticity is essential to ensure plasticity of biological processes in face of changing environmental conditions and, as such, its role in both the acceleration or impairment of evolutionary processes have been long debated [4]. Aside from being an essential component of all the developmental programs, plasticity underlies tumour development and progression, and it follows therefore that stochasticity may play an important role in a cancer-related context where chaos and unwonted interactions are predominant. However, the mechanisms leading to the emergence of patterns and conservation of biological functions from this noise remains elusive and the source of much debate. Pervasive transcription of the human cancer genome produces a multitude of lowly expressed long non-coding RNA (lncRNAs) often non-conserved and expressed at low copy number/cell [5] . Despite the above, these transcripts have been shown to impact key biological processes, such as cell viability, drug responses and/or tumour progression, by influencing multiple and highly expressed targets [6] . In this way they promote aberrant interactions and contribute to cancer interactome rewiring. Here we discuss recent findings in the field of RNA and condensate biology supporting a role for lncRNAs in cancer cell compartmentalization through regulation of aberrant phase separation. By way of this process, lncRNAs demonstrate their key role in converting stochastic signals arising from pervasive transcription into cellular organization, thus explaining their contribution to the aberrant cancer interactome (Table 1).

Cellular compartmentalization: order emerging from chaos

Biochemical oscillations in time and space control every aspect of cellular and organismal physiology (e.g. cell cycle, circadian rhythms, etc.) [7] and are

responsible for the formation of spatial patterns such as vertebrate segmentation and/or skin organization [8] . Conserved structural features (or compartments) in human cells and tissues and their perturbation (e.g. cellular and tissue patterns and recurrent mutations), have been used for centuries to classify diseases including cancer (e.g. nucleoli). Compartmentalization underlies the difference between eukaryotic and prokaryotic cells and provides spatiotemporal control over a number of cellular activities and biochemical reactions. While a handful of organelles are delimited by a lipid bilayer, many of them are membraneless (i.e. the P-bodies, the nucleolus, the nuclear speckles) and rely on a principle called phase separation (BOX1) or liquid unmixing for their formation [9] . Membraneless bodies phase-separate and become immiscible when their components reach the solubility limit [10] and their origin is therefore dependent on physical properties, such as temperature, concentration and pH. As such, they can adjust the rate of intracellular reactions in an environmentally-tunable fashion [11] and thus in response to biochemical oscillations. Essential for the nucleation of several membraneless bodies are specific combinations of RNAs [12, 13] and RNA-binding proteins containing Prion-Like Domains (PLDs) and/or **Intrinsically Disordered Regions (IDR)** which are generally prone to multivalent interactions. More than 30% of the eukaryotic proteome contains IDRs [14] and the motifs are particularly enriched in proteins with a role in cytoskeleton assembly and signal transduction [15], indicating that a larger number of molecular processes essential for (cancer) cell survival, differentiation and migration may be directly regulated by phase separation.

The process of phase separation and its consequences for gene expression have been better studied in the nucleus, but membraneless compartments exist also in the cytoplasm where they are essential in conveying environmental stress signals to the nucleus. Examples of membraneless bodies in the cytosol are Stress Granules (SG) and P-Bodies (PBs). (Box 2)

LncRNAs: the CC (concentration and conservation) paradox

RNA is considered a key molecule in all the theories of the origins of life and it is certainly the central player in the “RNA world hypothesis” that posits that 4 billion years ago life began on earth starting from primitive RNA molecules. The theory is based on the observation that, compared to DNA and proteins, RNA is a more flexible

molecule capable of storing the genetic information as does DNA (e.g. RNA viruses), but also possessing catalytic abilities like proteins (e.g. ribozymes). Additionally, RNA is notably highly responsive to environmental changes, riboswitches in prokaryotes for instance, can detect specific metabolites and modify their conformation to activate/inactivate gene expression in response to these changes [16]. It therefore not surprising that beside being a client, RNA can also actively participate to the enucleation of membraneless bodies [12].

It is now widely accepted that the human genome is pervasively transcribed into a plethora of highly processed and regulated transcripts that are mostly non-protein-coding [5]. The non-coding genome hosts the vast majority of recurrent somatic mutations [17], copy number alterations [18] and cancer-related SNPs [19] and encodes for a class of transcripts -longer than 200 nucleotides- called lncRNAs [20] Interestingly, while only a minority are widely expressed, evolutionary conserved such as NEAT1 [21] or MALAT1 [22] the vast majority of lncRNAs are primate-specific (80%) and display low and cancer-restricted expression (Fig.1) [20, 23] . Primate and (cancer) cell specificity have been linked to the enrichment in Transposable Elements (TE), which occupy almost half of the human genome, at lncRNA loci. TE contain cis-regulatory sequences that can act as promoter and enhancers [24] . In particular, TE belonging to the class of Endogenous RetroViruses (ERVs) [25] are enriched at the Transcriptional Start Site (TSS) of human lincRNA (long-intergenic non-coding RNAs) genes. TE are often methylated and kept silent, however hypomethylation and reactivation have been detected in cancer cells [26]; additionally somatic mutations and chromosomal rearrangement can also contribute to reactivation of these sequences [27] and thus to the evolution and emergence of lncRNA sequences [24] during the course of cancer. The lack of overt sequence conservation -and consequently of genetic models- has been used to undermine the importance of lncRNAs, however they still display important evolutionary conserved functions [28] (conservation paradox) which converge towards the regulation of epigenetic and post-transcriptional events [20]. One outstanding question in the lncRNA field is how these molecules, expressed at low copy number/cell, can impact key biological processes by influencing multiple and highly expressed targets. As already proposed in 2018 [6], emerging evidence suggests that the induction of phase separation may underlie a common mechanism exploited by lncRNAs to exert

their functions [29-31] . Indeed, 75% of human long non-coding transcripts contain at least a partial retroviral insertion and thus repetitive sequences [24] that may naturally act as molecular crowders to rewire cellular compartmentalization in response to environmental cues. Such a model may solve the conservation and concentration paradox by explaining how stochastic events such as the aberrant expression of lncRNAs can give rise to conserved functions even at low concentrations [32] . Supporting this hypothesis, it was recently demonstrated that two copies of Xist, a lncRNA necessary for X inactivation in placental mammals and implicated in the development of haematological cancers in *Mus musculus* [33], are sufficient to enucleate 50 macromolecular foci, containing the critical silencing protein SPEN [32]. Overall, these observations suggest that a re-evaluation of the concepts of specificity and stoichiometry in RNA-protein interactions, may be necessary. Furthermore, the consequences of the above on the potential of aberrant lncRNAs expression in cancer need to be considered [34] .

Cancer: apocalypse now

What happens when a lncRNA shows up in the cancer process?

Relaying on the interplay between individual genetic background, epigenetics and environmental factors, cancer development and progression is, by definition, a stochastic event [35] . The aberrant expression of a lncRNA in this context, may therefore eventually bring a new order to the chaos of the cancer cell by reshaping biochemical reactions (Fig. 2). The melanoma-specific lncRNA SAMMSON, for instance, coordinates to boost rRNA maturation and protein synthesis in the mitochondria and in the cytoplasm, by trapping the nuclear protein CARF in the cytoplasm in complex with the mitochondrial protein p32 [36]. Additionally paraspeckles, nuclear membraneless bodies assembled around the lncRNA NEAT1 [37], essential for cancer initiation and progression [20, 38, 39] can affect responses to therapy by sequestering an essential molecular complex like the Integrator Complex, necessary for the processing of the 3'-end of all RNAs [40].

Although these concepts have only recently gained traction in the literature, the role of RNA and more specifically of lncRNAs as scaffolds [41] and the role of aberrant membraneless compartmentalisation in cancer has been widely accepted for far longer. An Italian pathologist G. Pianese (in 1896) realized that shape and number of the nucleoli- nuclear structures that form upon phase separation following the

interaction of rRNAs and Alu B1-related RNAs with the intrinsically disordered proteins fibrillarin, nucleolin and nucleophosmin [42, 43] - is altered in carcinomas. Since this discovery, this parameter has been used as a cancer biomarker [44] .

It is arguably important to ask therefore, whether the role of molecular crowders such as lncRNAs in cancer phenotypes has been overlooked, by restriction of investigations to conserved lncRNAs with a well-established role in physiology, instead of looking for gain of function of aberrantly expressed molecules.

Concluding Remarks

The discovery of a regulatory role played by RNA in the biogenesis of condensates [45] and the established role of some of them in cancer, open new exciting possibilities not only for the specific targeting of these membraneless bodies, but also for a better understanding of biology of these somewhat enigmatic molecules. Whether phase separation is a common mechanism exploited by lncRNAs has yet to be determined, however this possibility would allow us to reconcile many difficult to explain findings. In this opinion article, we have summarized some unanswered questions in the field of lncRNAs and highlighted how some of these could be explained by a model involving molecular crowding and phase separation as the main mediator of lncRNA molecular functions. In this sense lncRNAs may convert the oscillatory signals coming from the tumour microenvironment into clear compartments and contribute to the organized chaos of cancer cells. As such lncRNAs would be at the crossroad of stochastic and **deterministic** events. The above considerations necessitate a full revaluation of the concepts of **specificity** and **stoichiometry** in RNA-protein interactions. To achieve this, the effect of physical changes in the Tumour MicroEnvironment (TME) on RNA compartmentalization and RNA-dependent protein interactions (transcriptome-wide) should be determined. Whether or not a correlation exists the identification and in-dept characterisation of the biology of specific lncRNAs implicated in LLPS would be an important step in our understanding of the basis of these interactions. Towards this, studies on the structure and modification of lncRNAs would therefore certainly be important, since the knowledge of the mechanism could then inform the design of synthetic RNAs that can promote LLPS and/or of cancer-specific inhibitors of selected membraneless bodies. Additionally, the patterns produced by specific lncRNAs could be used to

design sophisticated diagnostic test based on high-content imaging of cancer samples (see Outstanding Question box).

Acknowledgements

The authors declare no competing interests.

We would like to thank M. Leucci for manuscript proofreading. The cartoon in figure 3 was prepared using Somersault libraries available from www.somersault1824.com.

E.L. is funded by the Melanoma Research Alliance (MRA) young investigator https://doi.org/10.48050/pc.gr.80542, by Stichting Tegen Kanker (FAF-F/2018/1184) and by a KU Leuven C1 grant (no. #C16/19/006).

Box 1: Liquid-Liquid Phase separation (LLPS)

The term LLPS describes the spontaneous demixing of a homogenous solution into two or more phases when homotypic interactions are energetically favored over the entropic tendency of the solution to remain mixed. Demixing occurs when molecules in solution reach their solubility limit and the process is therefore obeying to polymer physics' principles rather than to classical stoichiometry rules.

Once the threshold for the formation of condensates has been reached under specific physico-chemical conditions (e.g critical concentration, temperature, salinity, pH and/or electrostatic and hydrophobic interactions) [46], the membraneless bodies establish a dynamic equilibrium with the surrounding environment allowing them to grow or dissolve without a net change in concentration. In a physiological context, LLPS leads to the formation of a dense phase, where proteins, DNA and RNA are 10-100 folds more concentrated than the surrounding dilute phase[47].

The molecules initiating LLPS are called scaffolds as they are necessary for the formation of specific condensates. The molecules recruited to the condensate but not necessarily to engage LLPS are known as clients [13]. For instance, under specific stress conditions, stalled PreInitiation Complex (PIC) mRNPs and the two RNA Binding Proteins (RBPs) Ras-GTPase-activating protein SH3-domain-Binding-Protein 1 (G3BP1) and T-cell-restricted Intracellular Antigen-1(TIA-1) are crucial nucleators triggering LLPS of stress granules[48-50]. Other proteins with various functions are subsequently recruited to the core, allowing dynamic RNA-protein exchange, such as Caprin1 and Ubiquitin Specific Peptidase 10 (USP10), two G3BP competitive binders that promote or inhibit SG condensation, respectively [51-53].

LLPS relies on weak cooperative interactions, and/or strong interactions reversible in a short timescale [54]. Therefore, poorly structured biomolecules and those containing repetitive elements, such has IDRs, are more prone to phase separation [55]. As such, arginine-and glycine-rich (RGG/RG) repeats and PLDs are two important classes of stickers found in proteins driving biocondensates' assembly [56, 57].

Additionally, RNA which is a negatively charged molecule, containing posttranscriptional modifications, often provides stickers for the binding of multiple RBPs making it a key regulator of condensates' formation, properties and dynamics.

Box 2: Cytoplasmic RNA Granules

RNA granules are membraneless condensates composed of protein-enriched RNA species that contribute to all steps of RNA metabolism namely: processing; transport; storage; translation and/or degradation [58]. Among them, PBs and SGs are to date the most well studied mRNA silencing *foci*[53].

PBs and SGs are cytoplasmic foci ranging in size from 400 to 500nm[59] and 100 to 2000nm[50], respectively. Differently from SGs that arise upon exposure to stress, PBs are constitutive but their size and number increase under stress [60, 61].

SGs' assembly requires two steps. First, the inhibition of translation initiation either through phosphorylation of eukaryotic initiation factor 2 alpha (eIF2), through mTOR inhibition, or through interference with the eIF4F complex, all of which lead to the disassembly of polysomes[62, 63]. Secondly, the condensation of stalled pre-initiation complexes and their associated RBPs into distinct phase-separated granules[63] regulated by a variety of proteins such as G3BP, PolyA-Binding Protein (PABP) and TIA-1 called SG nucleators for their ability to nucleate SGs in the absence of stress when overexpressed *in vitro*[49, 50, 60]. All of them are characterized by their IDRs and RNA-Binding Domains (RBDs) favoring multivalent interactions and thus macromolecular aggregates formation[63]. Similarly to SGs and nuclear condensates, PBs rely on complex RNA – protein interactions, IDR-enriched protein sequences and LLPS for their formation[42, 60, 63].

In line with their assembly process, SGs enclose mainly proteins associated with translation initiation such as the 40S ribosomal subunit, PABP and eiF4G1[60, 61] and the RBP G3BP1 with its two crucial partners Caprin1 and USP10 [64]. PBs are, on the other hand, predominantly composed of mRNA decay proteins such as the deadenylation complex CNOT1, mRNA DeCaPping enzyme subunits 1a and 2 (DCP1a, DCP2) and decapping activators such as EDC3 and EDC4[60]. Although different from one another, SGs and PBs share many common proteins such as TIA-1, FASTK including those promoting association between both granules such as TTP, BRF1 and eIF4E[60, 61].

In general, both SGs and PBs assemble and disassemble rapidly (within minutes) upon stress induction and removal respectively[60]. Despite this, different types of environmental stress (amino acid starvation, UV irradiation, oxidative and/or osmotic stress, ER stress, etc.) can lead to distinct RNA granules subtypes discernible by their protein composition and dynamics[63, 65]. To date, at least three different subtypes of SGs have been identified[53, 60, 63]: type I SGs form upon stress-

induced phosphorylation of eiF2α (e.g. oxidative stress, ER stress and viral infection) and require G3BP and 48S PICs for their assembly; type II SGs still require G3BP however, they form independently from eiF2α phosphorylation; type III SGs lack eiF3, differently from the other two subtypes, and their assembly is triggered by UV, glucose and starvation, nitric oxide and other chemical compounds [63].

While SGs assembly is unnecessary for translational repression during stress, it may enhance the translational rewiring process[63, 66] by segregating translationally stalled mRNAs, that mostly encode for housekeeping proteins such as GAPDH and B-Actin [63, 67, 68]. Likewise, PBs are not the sites of mRNA degradation, but rather for the storage of repressed mRNAs awaiting either translation or decay[60, 65].

Formation of PBs and SGs in tumor cells in response of stress is important for adaptation [69]. In keeping with this, SGs assembly has been significantly observed in many cancer types (e.g., pancreatic cancer, glioblastoma) and often associated with drug tolerance[69, 70].

References

1 Chen, J., *et al.* (2020) Pervasive functional translation of noncanonical human open reading frames. *Science* 367, 1140-1146

2 Eling, N., *et al.* (2019) Challenges in measuring and understanding biological noise. *Nature reviews. Genetics* 20, 536-548

3 Zhang, Y., *et al.* (2021) Random heterogeneity outperforms design in network synchronization. *Proceedings of the National Academy of Sciences of the United States of America* 118

4 Matic, I. (2019) Mutation Rate Heterogeneity Increases Odds of Survival in Unpredictable Environments. *Molecular cell* 75, 421-425

5 Carninci, P., *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science* 309, 1559-1563

6 Leucci, E. (2018) Long Noncoding RNA as Novel Cancer Diagnostic and Effective Therapeutic Targets. In *Applied RNA Bioscience*, pp. 189-202

7 Novák, B. and Tyson, J.J. (2008) Design principles of biochemical oscillators. *Nature reviews. Molecular cell biology* 9, 981-991

8 Quiroz, F.G., *et al.* (2020) Liquid-liquid phase separation drives skin barrier formation. *Science* 367

9 Boeynaems, S., *et al.* (2018) Protein Phase Separation: A New Phase in Cell Biology. *Trends in cell biology* 28, 420-435

10 Banani, S.F., *et al.* (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nature reviews. Molecular cell biology* 18, 285-298

11 Gomes, E. and Shorter, J. (2019) The molecular language of membraneless organelles. *The Journal of biological chemistry* 294, 7115-7127

12 Polymenidou, M. (2018) The RNA face of phase separation. *Science* 360, 859-860

13 Roden, C. and Gladfelter, A.S. (2021) RNA contributions to the form and function of biomolecular condensates. *Nature reviews. Molecular cell biology* 22, 183-195

14 Bergeron-Sandoval, L.P., *et al.* (2016) Mechanisms and Consequences of Macromolecular Phase Separation. *Cell* 165, 1067-1079

15 Miao, Y., *et al.* (2018) Phospho-regulation of intrinsically disordered proteins for actin assembly and endocytosis. *The FEBS journal* 285, 2762-2784

16 Ray, S., *et al.* (2019) Kinetics coming into focus: single-molecule microscopy of riboswitch dynamics. *RNA biology* 16, 1077-1085

17 Melton, C., *et al.* (2015) Recurrent somatic mutations in regulatory regions of human cancer genomes. *Nature genetics* 47, 710-716

18 Beroukhim, R., *et al.* (2010) The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899-905

19 Cheetham, S.W., *et al.* (2013) Long noncoding RNAs and the genetics of cancer. *British journal of cancer* 108, 2419-2425

20 Leucci, E. (2018) Cancer development and therapy resistance: spotlights on the dark side of the genome. *Pharmacology & therapeutics* 189, 22-30

21 Adriaens, C., *et al.* (2016) p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nature medicine* 22, 861-868

22 Arun, G., *et al.* (2020) MALAT1 Long Non-Coding RNA: Functional Implications. *Non-coding RNA* 6

23 Leucci, E., *et al.* (2016) Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* 531, 518-522

24 Ganesh, S. and Svoboda, P. (2016) Retrotransposon-associated long non-coding RNAs in mice and men. *Pflugers Archiv : European journal of physiology* 468, 1049-1060

25 Kelley, D. and Rinn, J. (2012) Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome biology* 13, R107

26 Szpakowski, S., *et al.* (2009) Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements. *Gene* 448, 151-167

27 Rodriguez-Martin, B., *et al.* (2020) Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nature genetics* 52, 306-319

28 Ulitsky, I., *et al.* (2011) Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147, 1537-1550

29 Daneshvar, K., *et al.* (2020) lncRNA DIGIT and BRD3 protein form phase-separated condensates to regulate endoderm differentiation. *Nature cell biology* 22, 1211-1222

30 Elguindy, M.M. and Mendell, J.T. (2021) NORAD-induced Pumilio phase separation is required for genome stability. *Nature* 595, 303-308

31 Quinodoz, S.A., *et al.* (2021) RNA promotes the formation of spatial compartments in the nucleus. *Cell* 184, 5775-5790.e5730

32 Markaki, Y., *et al.* (2021) Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* 184, 6174-6192.e6132

437 33 Yildirim, E., *et al.* (2013) Xist RNA is a potent suppressor of hematologic cancer in mice.
438 *Cell* 152, 727-742
439 34 Carlevaro-Fita, J., *et al.* (2020) Cancer LncRNA Census reveals evidence for deep
440 functional conservation of long noncoding RNAs in tumorigenesis. *Communications biology*
441 3, 56
442 35 Tomasetti, C. and Vogelstein, B. (2015) Cancer etiology. Variation in cancer risk among
443 tissues can be explained by the number of stem cell divisions. *Science* 347, 78-81
444 36 Vendramin, R., *et al.* (2018) SAMMSON fosters cancer cell fitness by concertedly
445 enhancing mitochondrial and cytosolic translation. *Nature structural & molecular biology*
446 25, 1035-1046
447 37 Fox, A.H., *et al.* (2018) Paraspeckles: Where Long Noncoding RNA Meets Phase
448 Separation. *Trends in biochemical sciences* 43, 124-135
449 38 Mello, S.S., *et al.* (2017) Neat1 is a p53-inducible lincRNA essential for transformation
450 suppression. *Genes & development* 31, 1095-1108
451 39 Naveed, A., *et al.* (2021) NEAT1 polyA-modulating antisense oligonucleotides reveal
452 opposing functions for both long non-coding RNA isoforms in neuroblastoma. *Cellular and*
453 *molecular life sciences : CMLS* 78, 2213-2230
454 40 Barra, J., *et al.* (2020) Integrator restrains paraspeckles assembly by promoting isoform
455 switching of the lncRNA NEAT1. *Science advances* 6, eaaz9072
456 41 Wang, K.C. and Chang, H.Y. (2011) Molecular mechanisms of long noncoding RNAs.
457 *Molecular cell* 43, 904-914
458 42 Feric, M., *et al.* (2016) Coexisting Liquid Phases Underlie Nucleolar Subcompartments.
459 *Cell* 165, 1686-1697
460 43 Verheyden, Y., *et al.* (2018) Control of nucleolar stress and translational reprogramming
461 by lncRNAs. *Cell stress* 3, 19-26
462 44 Stamatopoulou, V., *et al.* (2018) Use of the iNo score to discriminate normal from altered
463 nucleolar morphology, with applications in basic cell biology and potential in human disease
464 diagnostics. *Nature protocols* 13, 2387-2406
465 45 Maharana, S., *et al.* (2018) RNA buffers the phase separation behavior of prion-like RNA
466 binding proteins. *Science* 360, 918-921
467 46 Alberti, S., *et al.* (2019) Considerations and Challenges in Studying Liquid-Liquid Phase
468 Separation and Biomolecular Condensates. *Cell* 176, 419-434
469 47 Li, P., *et al.* (2012) Phase transitions in the assembly of multivalent signalling proteins.
470 *Nature* 483, 336-340

471 48 Yang, P.*, et al.* (2020) G3BP1 Is a Tunable Switch that Triggers Phase Separation to
472 Assemble Stress Granules. *Cell* 181, 325-345.e328
473 49 Tourrière, H.*, et al.* (2003) The RasGAP-associated endoribonuclease G3BP assembles
474 stress granules. *The Journal of cell biology* 160, 823-831
475 50 Gilks, N.*, et al.* (2004) Stress granule assembly is mediated by prion-like aggregation of
476 TIA-1. *Molecular biology of the cell* 15, 5383-5398
477 51 Solomon, S.*, et al.* (2007) Distinct structural features of caprin-1 mediate its interaction
478 with G3BP-1 and its induction of phosphorylation of eukaryotic translation initiation factor
479 2alpha, entry to cytoplasmic stress granules, and selective interaction with a subset of
480 mRNAs. *Molecular and cellular biology* 27, 2324-2342
481 52 Campos-Melo, D.*, et al.* (2021) The Integral Role of RNA in Stress Granule Formation
482 and Function. *Frontiers in cell and developmental biology* 9, 621779
483 53 Ivanov, P.*, et al.* (2019) Stress Granules and Processing Bodies in Translational Control.
484 *Cold Spring Harbor perspectives in biology* 11
485 54 Alberti, S. and Dormann, D. (2019) Liquid-Liquid Phase Separation in Disease. *Annual*
486 *review of genetics* 53, 171-194
487 55 Dignon, G.L.*, et al.* (2020) Biomolecular Phase Separation: From Molecular Driving
488 Forces to Macroscopic Properties. *Annual review of physical chemistry* 71, 53-75
489 56 Chong, P.A.*, et al.* (2018) RGG/RG Motif Regions in RNA Binding and Phase Separation.
490 *Journal of molecular biology* 430, 4650-4665
491 57 Alberti, S.*, et al.* (2009) A systematic survey identifies prions and illuminates sequence
492 features of prionogenic proteins. *Cell* 137, 146-158
493 58 Moujaber, O. and Stochaj, U. (2018) Cytoplasmic RNA Granules in Somatic Maintenance.
494 *Gerontology* 64, 485-494
495 59 Ayache, J.*, et al.* (2015) P-body assembly requires DDX6 repression complexes rather
496 than decay or Ataxin2/2L complexes. *Molecular biology of the cell* 26, 2579-2595
497 60 Riggs, C.L.*, et al.* (2020) Mammalian stress granules and P bodies at a glance. *Journal of*
498 *cell science* 133
499 61 Kedersha, N.*, et al.* (2005) Stress granules and processing bodies are dynamically linked
500 sites of mRNP remodeling. *The Journal of cell biology* 169, 871-884
501 62 Panas, M.D.*, et al.* (2016) Mechanistic insights into mammalian stress granule dynamics.
502 *The Journal of cell biology* 215, 313-323
503 63 Hofmann, S.*, et al.* (2021) Molecular mechanisms of stress granule assembly and
504 disassembly. *Biochimica et biophysica acta. Molecular cell research* 1868, 118876

64 Kedersha, N., *et al.* (2016) G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *The Journal of cell biology* 212, 845-860
65 Tian, S., *et al.* (2020) RNA Granules: A View from the RNA Perspective. *Molecules (Basel, Switzerland)* 25
66 Advani, V.M. and Ivanov, P. (2020) Stress granule subtypes: an emerging link to neurodegeneration. *Cellular and molecular life sciences : CMLS* 77, 4827-4845
67 Anderson, P. and Kedersha, N. (2009) Stress granules. *Current biology : CB* 19, R397-398
68 Kedersha, N. and Anderson, P. (2002) Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochemical Society transactions* 30, 963-969
69 Gao, X., *et al.* (2019) Stress granule: A promising target for cancer treatment. *British journal of pharmacology* 176, 4421-4433
70 Vilas-Boas Fde, A., *et al.* (2016) Impairment of stress granule assembly via inhibition of the eIF2alpha phosphorylation sensitizes glioma cells to chemotherapeutic agents. *Journal of neuro-oncology* 127, 253-260
71 Lee, B.T., *et al.* (2021) The UCSC Genome Browser database: 2022 update. *Nucleic acids research*
72 Verfaillie, A., *et al.* (2015) Decoding the regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. *Nature communications* 6, 6683

Glossary (500 words)

Deterministic: the term refers to events that develop according to a plan (non-random) and thus are predictable.

IDR: also called Intrinsically Disordered Regions, are domains in proteins that do not contain a defined 3D structure in physiological conditions. They are often found at flexible linkers and loops connecting different domains. IDRs contains amino acids with high net charge and low hydrophobicity.

Pervasive Transcription: the term refers to the finding that in most species the genome is almost entirely transcribed including area before considered as purely regulatory.

Specificity: the property of a certain molecules to interact with selected partners. Specificity is often conferred by complementary 3D structural or sequence motifs.

Stochastic: the term refers to a process fitting a random distribution and thus lacking a plan.

Stoichiometry: is the numerical relationship between reactants and products in a chemical reaction.

Figure legends

Figure 1: The Conservation & Concentration paradox: lncRNA *SAMMSON* as a paradigm. The lncRNA SAMMSON is hosted downstream of the protein coding gene MITF. In contrast with MITF, SAMMSON is primate-specific (like 80% of lncRNAs) and lowly expressed (5<x<200 copies/cell) in melanoma lines [71, 72]. Furthermore, SAMMSON is enriched in repetitive sequences. Table 1 reports key well characterized lncRNAs; in blue transcripts phase separating or localizing at membraneless bodies.

Figure 2: Potential role of lncRNA-induced phase separation in Cancer

The aberrant expression of lncRNAs is induced by extracellular cues during tumour development and progression. These lncRNAs act as molecular crowders and thus regulate phase separation to induce specific cell states and cancer phenotypes in response to changes in the physical properties of the tumour microenvironment. Examples of lncRNAs implicated in cancer and/or known to phase separate have been highlighted [23, 30, 32].

Table 1

lncRNA	Localization	GTEx expression (average TPM)	Structure Features	Interactors	Function	Physiopathological process	References[a]
MALAT1	Nucleus	826.9 (ovaries)	tRNA-like small RNA at its 3'end	Transcription factors,Splicing factors, Epigenetic regulators	Regulates the phosphorylation of SR proteins in nuclear speckles, thus, modulates pre-mRNA splicing	Upregulated in many human malignancies. Correlates with poor prognosis and metastasis.	[S1-S4] [22]
NEAT1	Nucleus	671.2 (thyroid)	Repetitive RNA subdomains (long isoform); tRNA-like small RNA at its 3'end (long isoform)	Paraspeckle components (e.g. CARM1, FUS, p54nrb, PSPC1)	Drives LLPS of paraspeckles involved in gene expression regulation	Essential for skin cancer initiation and progression	[21] [S5, S6]
Xist	Nucleus	148.4 (ovaries)	Repetitive RNA subdomains (A-repeats and C-repeats)	Chromatin remodeling factors (e.g. Spen, Rbm15, Wtap)	Mediates the X-chromosome inactivation process by enriching repressive complexes to chromatin, possibly through LLPS in mouse	Upregulated in colorectal cancer and correlates with poor overall survival.	[S7, S8]
HOTAIR	Nucleus	27.1 (arteries)	Not known	PRC2, LSD1	Mediates transcriptional repression OF *HOXD* gene independently of PRC2	Highly expressed and involved in initiation and progression of different cancers (e.g. breast cancer)	[S9-S11]

SAMMSON	Nucleus & Cytoplasm	1 (arteries)	Not known	CARF, p32	In melanoma: Favors an aberrant interaction between p32 and CARF in the cytosol and sequesters CARF away from its partner XRN2 resulting in an increase in ribosome biogenesis.	Upregulated in melanoma and promotes tumor growth. Its knockdown increases the response of melanoma patient-derived xenografts to targeted therapy.	[23, 36]
TINCR	Cytoplasm	108.5 (skin)	Not known	STAU1	Binds STAU1 to mediate stabilization epidermal differentiation mRNAs (e.g. *KRT80*). In melanoma, interacts with pro-invasive RNAs such as *ATF4,* inhibiting their binding to ribosomes and the acquisition of invasive phenotype.	Aberrantly expressed in many cancers. Exerts both tumor-suppressive and oncogenic effects, therefore modulating cancer progression.	[S12-S14]
AGPG	Nucleus and Cytoplasm	Not known	Not known	PFKFB3	Binds and stabilizes PFKFB3 promoting its enrichment in cancer cells, leading to enhanced glycolytic flux and cell cycle progression	Upregulated in many cancers and is associated with poor prognosis (e.g. esophageal squamous cell carcinoma). Its depletion impedes tumor growth in PDX models.	[S15]
NORAD	Cytoplasm	285 (Brain)	Enriched with pumilio response elements (PREs)	Pumilio proteins (PUM1, PUM2)	Inhibits the activity of PUM proteins via nucleation of PUM condensates (NP bodies) to promote genomic stability	Aberrantly expressed in various cancers and involved in carcinogenesis processes (e.g. proliferation, invasion, metastasis, apoptosis)	[30] [S16]

DIGIT	Nucleus	Not known	Not known	BRD3	Promotes phase separation of BRD3 condensates and their recruitment to H3K18ac regions, thus regulating transcription factors of endoderm differentiation.	Not known	[29]
LINC-PINT	Nucleus	33.7 (ovaries)	Two highly conserved short regions	PRC2	Represses genes responsible for cancer cell invasion through its interaction with PRC2	Downregulated in multiple cancers	[S17]
LASTR	Nucleus	Not known	Not known	RNA-splicing factor SART3	Promotes splicing efficiency by regulating SART3 binding to the U4 and U6 snRNPs	Essential for the growth of triple negative breast tumors	[S18]
DilncRNA	Nucleus	Not known	Not known	DNA Damage Response (DDR) RNAs and proteins (e.g. 53BP1)	Synthetized at DNA Double Strand Breaks (DSB), it interacts with DDR proteins to promote phase separation of DDR foci responsible for transcriptional regulation.	DDR dysfunction has been reported in a plethora of human malignancies	[S19, S20]
PNCTR	Nucleus	Not known	Enriched with Short-Tandem Repeats (STRs)	PTBP1	Sequesters PTBP1 to form a phase-separated body named peri-nucleolar compartment (PNC) thus inhibiting PTBP1 splicing activity and promoting cell survival	Highly expressed in a multitude of cancers.	[S21]

HSATIII	Nucleus	Not known	Enriches with STRs (GGAAU)n	HNRNPs, STLM, NCOA5, SAFB	Acts as a scaffold for the formation of nuclear stress bodies upon thermal stress, regulating gene expression	Not known	[S22, S23]
TNBL	Nucleus	Not known	Derived from NBL2 repeats	NPM1, SAM68 and CELF1	Upon NBL2 DNA hypomethylation and histone acetylation in colorectal cancer, TNBL expression is increased leading to the formation of aggregates close to NBL2 loci where it interacts with SAM68 involved in splicing regulation	TNBL-SAM68 perinucleolar bodies are cancer-specific aggregates. Their role is still to be determined.	[S24]

[a] See the supplemental information online.

Outstanding questions

1. Is phase separation a general mechanism exploited by lncRNAs to exert their functions? If so, can we identify the stickers driving LLPS of lncRNAs?
2. Could specific dynamics of lncRNA driven phase separation be used as markers of disease?
3. Can we target specific membraneless compartments by targeting the corresponding RNA?
4. Can we engineer specific compartments by using lncRNA modules?

Supplementary references:

S1. Wilusz, J.E., et al. (2008) 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell 135, 919-932
S2. Tripathi, V., et al. (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell 39, 925-938
S3. Li, B., et al. (2014) Activation of LTBP3 gene by a long noncoding RNA (lncRNA) MALAT1 transcript in mesenchymal stem cells from multiple myeloma. The Journal of biological chemistry 289, 29365-29375
S4. Zhou, L., et al. (2018) Long non-coding RNA MALAT1 interacts with transcription factor Foxo1 to regulate SIRT1 transcription in high glucose-induced HK-2 cells injury. Biochemical and biophysical research communications 503, 849-855
S5. Yamazaki, T., et al. (2018) Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. Molecular cell 70, 1038-1053.e1037
S6. Sunwoo, H., et al. (2009) MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome research 19, 347-359
S7. Moindrot, B., et al. (2015) A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. Cell reports 12, 562-572
S8. Cerase, A., et al. (2019) Phase separation drives X-chromosome inactivation: a hypothesis. Nature structural & molecular biology 26, 331-334
S9. Rinn, J.L., et al. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311-1323
S10. Portoso, M., et al. (2017) PRC2 is dispensable for HOTAIR-mediated transcriptional repression. The EMBO journal 36, 981-994
S11. Gupta, R.A., et al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464, 1071-1076
S12. Kretz, M., et al. (2013) Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature 493, 231-235
S13. Sharma, U., et al. (2020) Long non-coding RNA TINCR as potential biomarker and therapeutic target for cancer. Life sciences 257, 118035
S14. Melixetian, M., et al. (2021) Long non-coding RNA TINCR suppresses metastatic melanoma dissemination by preventing ATF4 translation. EMBO reports 22, e50852
S15. Liu, J., et al. (2020) Long noncoding RNA AGPG regulates PFKFB3-mediated tumor glycolytic reprogramming. Nature communications 11, 1507
S16. Yang, Z., et al. (2019) Noncoding RNA activated by DNA damage (NORAD): Biologic function and mechanisms in human cancers. Clinica chimica acta; international journal of clinical chemistry 489, 5-9
S17. Marín-Béjar, O., et al. (2017) The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. Genome biology 18, 202
S18. De Troyer, L., et al. (2020) Stress-induced lncRNA LASTR fosters cancer cell fitness by regulating the activity of the U4/U6 recycling factor SART3. Nucleic acids research 48, 2502-2517
S19. Michelini, F., et al. (2017) Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. Nature cell biology 19, 1400-1411

S20. Pessina, F., et al. (2019) Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. Nature cell biology 21, 1286-1299
S21. Yap, K., et al. (2018) A Short Tandem Repeat-Enriched RNA Assembles a Nuclear Compartment to Control Alternative Splicing and Promote Cell Survival. Molecular cell 72, 525-540.e513
S22. Valgardsdottir, R., et al. (2005) Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. Molecular biology of the cell 16, 2597-2604
S23. Aly, M.K., et al. (2019) Two distinct nuclear stress bodies containing different sets of RNA-binding proteins are formed with HSATIII architectural noncoding RNAs upon thermal stress exposure. Biochemical and biophysical research communications 516, 419-423
S24. Dumbovic, G., et al. (2018) A novel long non-coding RNA from NBL2 pericentromeric macrosatellite forms a perinucleolar aggregate structure in colon cancer. Nucleic acids research 46, 5504-5524

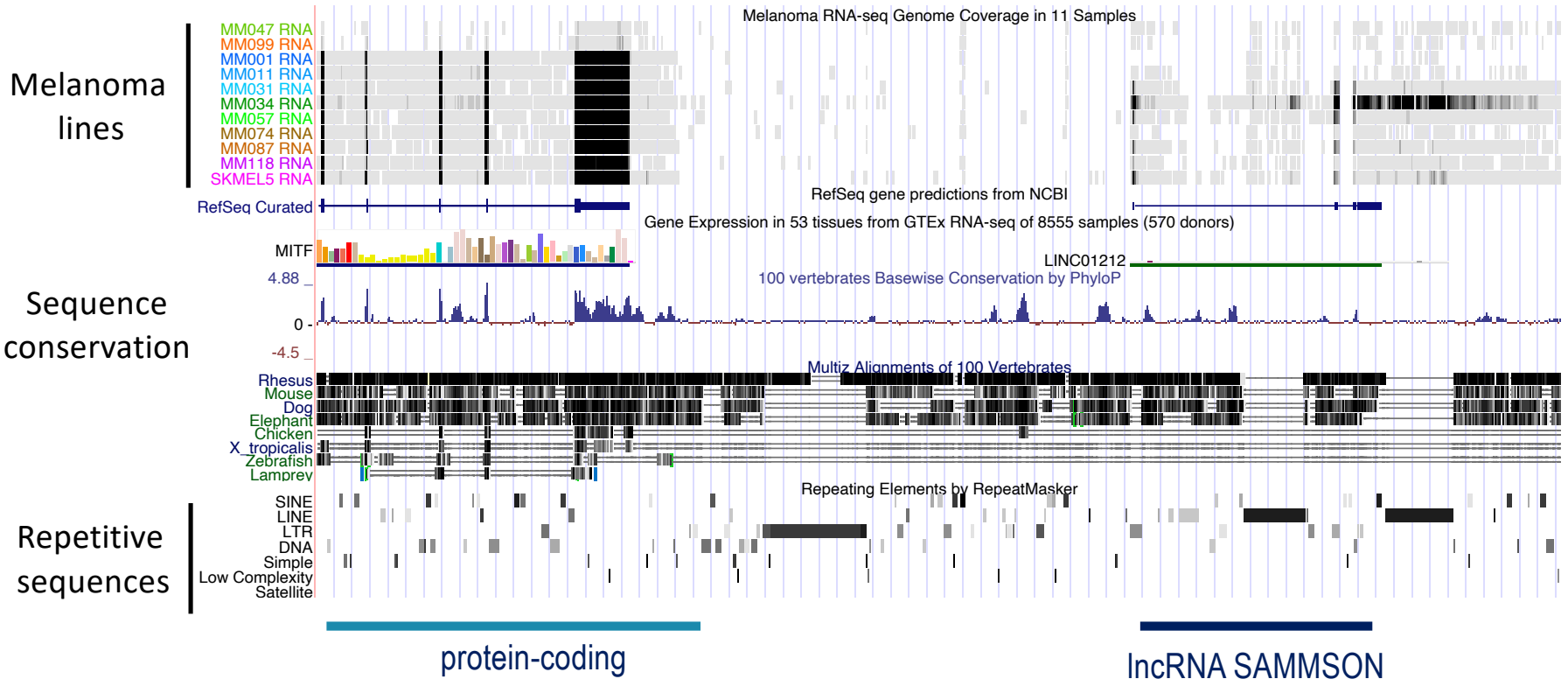
Melanoma lines
Sequence conservation
Repetitive sequences
Melanoma RNA-seq Genome Coverage in 11 Samples
MM047 RNA
MM099 RNA
MM001 RNA
MM011 RNA
MM031 RNA
MM034 RNA
MM057 RNA
MM074 RNA
MM087 RNA
MM118 RNA
SKMEL5 RNA
RefSeq Curated
RefSeq gene predictions from NCBI
Gene Expression in 53 tissues from GTEx RNA-seq of 8555 samples (570 donors)
MITF
LINC01212
4.88
0
-4.5
100 vertebrates Basewise Conservation by PhyloP
Multiz Alignments of 100 Vertebrates
Rhesus
Mouse
Dog
Elephant
Chicken
X_tropicalis
Zebrafish
Lamprey
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
protein-coding
lncRNA SAMMSON

Figure 1

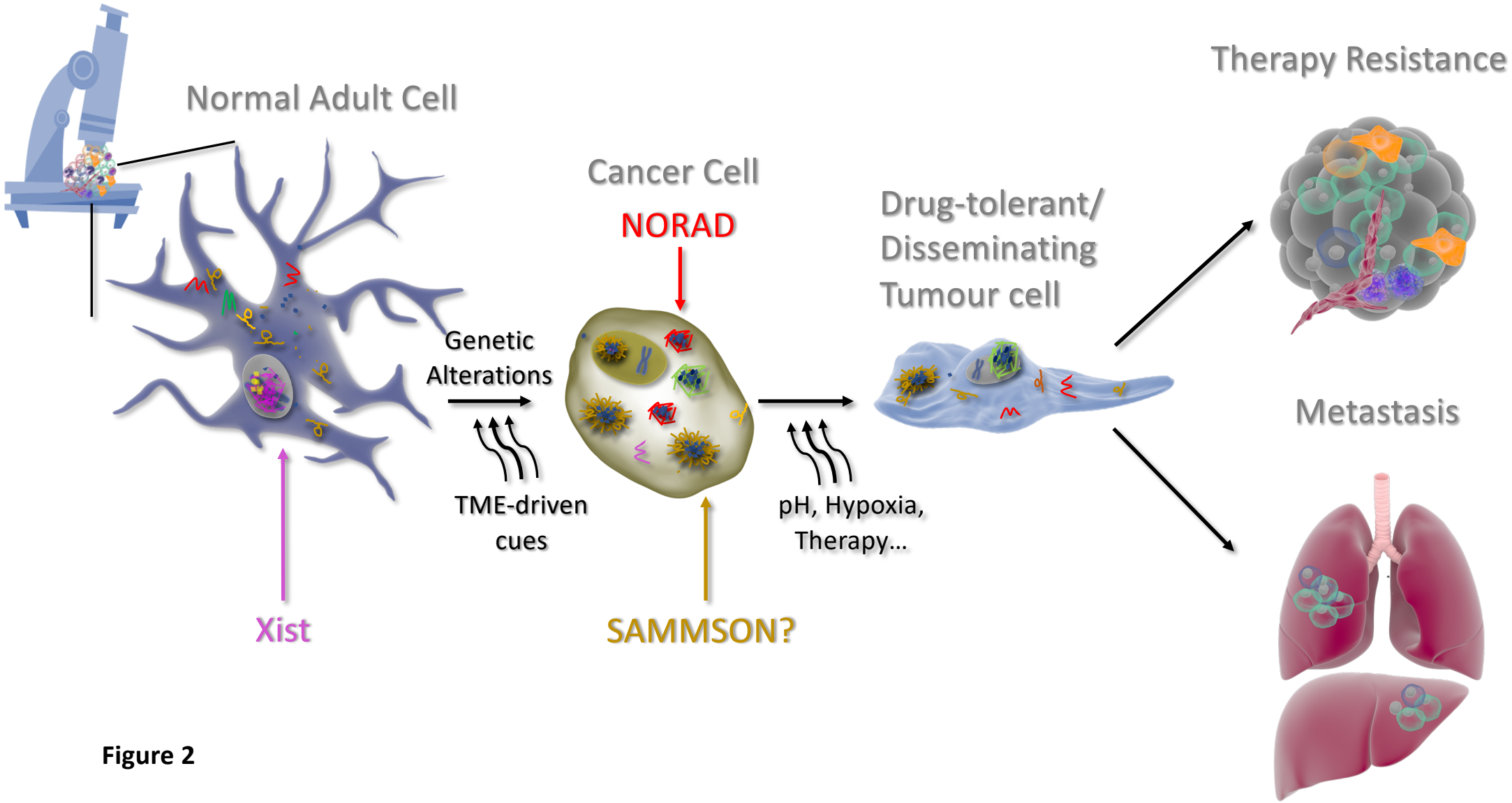
Normal Adult Cell
Cancer Cell
NORAD
Drug-tolerant/
Disseminating
Tumour cell
Therapy Resistance
Metastasis
Genetic
Alterations
TME-driven
cues
pH, Hypoxia,
Therapy…
Xist
SAMMSON?

Figure 2