Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
Leukemia. Author manuscript; available in PMC 2018 December 12.

Published in final edited form as:
Leukemia. 2018 December ; 32(12): 2626–2635. doi:10.1038/s41375-018-0116-y.

Long Intergenic Non-Coding RNAs have an Independent Impact on Survival in Multiple Myeloma

Mehmet Kemal Samur[1], **Stephane Minvielle**[2], **Annamaria Gulla**[1], **Mariateresa Fulciniti**[1], **Alice Cleynen**[3], **Anil Aktas Samur**[1], **Raphael Szalat**[1], **Masood Shammas**[1], **Florence Magrangeas**[2], **Yu-Tzu Tai**[1], **Daniel Auclair**[5], **Jonathan Keats**[6], **Paul Richardson**[1], **Michel Attal**[4], **Philippe Moreau**[7], **Kenneth C. Anderson**[1], **Giovanni Parmigiani**[1], **Hervé Avet-Loiseau**[4,#], and **Nikhil C. Munshi**[1,8,#]

[1]Dana Farber Cancer Institute and Harvard Medical School, Boston, MA, USA 02215

[2]Center for Cancer Research Nantes-Angers, UMR 892 Inserm-6299 CNRS-University of Nantes, IRS-UN, Nantes 4407, France

[3]Centre National de la Recherche Scientifique, and Montpellier University, Montpellier, France

[4]L'Institut Universitaire du CancerOncopole, Toulouse

[5]Multiple Myeloma Research Foundation, Norwalk, CT 06851

[6]The Translational Genomics Research Institute (TGen), Phoenix, AZ 85004

[7]Department of Hematology, University Hospital, Nantes 44093, France

[8]Veterans Administration Boston Healthcare System, West Roxbury, MA

Abstract

Although long intergenic non-coding RNAs(lincRNA) role in various cancers is described, their significance in Multiple Myeloma(MM) remains poorly defined. Here we have studied the lincRNA profile and their clinical impact in MM. We performed RNA-seq on MM cells from 308 newly-diagnosed and uniformly-treated patients, 16 normal plasma cells and utilized RNA-seq data from 532 newly-diagnosed patients from CoMMpass study to analyze for lincRNAs. We observed 869 differentially expressed lincRNAs in MM compared to normal plasma cells. We identified 14 lincRNAs associated with PFS and calculated a risk score to stratify patients. The median PFS between high vs low-risk groups was 17 months vs not-reached(NR); and OS 30 months vs NR, respectively($p<0.0001$ for both). In the independent validation dataset between high and low-risk groups, PFS was 27 vs 42 months(HR 2.06[1.44–2.96];$p<0.0005$); and 4-year OS 62% vs 86%(HR 2.76[1.51–5.05];$p<.0005$) confirming significant clinical relevance of lincRNA in MM. Importantly, lincRNA signature was able to further identify patients with

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

[#]**Corresponding Authors**: To whom correspondence should be addressed: Dr Nikhil C. Munshi, Dana-Farber Cancer Institute, 450 Brookline Avenue, Dana B106, Boston, MA 02215, USA, Phone: +1-617-632-4218, Fax +1-617-582-8608, nikhil_munshi@dfci.harvard.edu; Dr Hervé Avet-Loiseau, Centre de Recherches en Cancérologie de Toulouse, Phone: +33(0) 5 82 74 15 75, Fax: +33(0) 5 82 74 16 85, avetloiseau.herve@iuct-oncopole.fr.

CONFLICT OF INTEREST

There are no competing financial interests in relation to the work described.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

significant differential outcomes within each low and high-risk categories identified using standard risk categorization including cytogenetic/FISH, ISS, and MRD negative or positive. Our results suggest that lincRNAs have an independent effect on MM outcome and provide a rationale to evaluate its molecular and biological impact.

Introduction

In addition to acting as a messenger between DNA and protein[1], RNA also contains diverse sets of regulatory functions. Recent analysis of RNA repertoire has identified large numbers of non-coding transcripts, including long intergenic non-coding RNA (lincRNA), which have transcripts longer than 200 nucleotides, and are located are located between the protein coding genes. lincRNAs have been considered to provide regulatory functions, however, their precise role in cellular biology remains unclear [2, 3].

Many of these non-coding elements have tissue-specific expressions, which are controlled in a regulated manner, in correlation with distinct gene sets that influence critical biological roles, including cell cycle regulation, survival, and immune response[4]. Some of these non-coding elements are regulated by tumor suppressor genes or oncogenes such as cMYC[5] or have been reported directly as a tumor suppressor or oncogene[6]. Consequently, they are found to be expressed differentially in tumors, and have been linked to the transformation of healthy cells into tumor cells[4]. lincRNAs have been described to play an intermediary role in modulating transcriptome as well as affecting miRNA activity[7, 8]. Recently, a role of lincRNA has been described in myeloma with multiple functions, such as cell proliferation and apoptosis[9, 10], interactions with miRNA[11, 12], protein coding genes[13] and individual lincRNA relationship to MM progression[14, 15]. Many types of lincRNAs have also been proven to strongly influence prognosis in diverse array of cancers[16–18]. Over the years, numerous studies have described the impact of gene expression profile (GEP) on clinical outcomes; however, an integrative analysis, which incorporates more than one genomic correlate, is lacking. Using whole transcriptome sequencing, we can now study these regulatory elements and their precise role in clinical outcomes and disease progression. In the current study, we examined lincRNAs using RNAseq data to investigate their correlation with clinical outcome in MM.

METHODS

Patient samples

We sequenced CD138+ MM cells from 308 newly-diagnosed multiple myeloma patients from IFM-DFCI 2009 clinical trial (ClinicalTrials.gov Identifier: NCT01191060)[19] and 16 normal donor plasma cells. In this study, transplant-eligible newly diagnosed patients younger than 66 years of age were randomized to either receive 8 courses of the RVD (Revlimid®-Velcade®-Dexamethasone) regimen, comprising a conventional-dose therapy, or an intensive approach, with 3 RVD courses, followed by single high-dose melphalan (200 mg/m^2) with autologous stem cell support, and 2 additional RVD cycles as consolidation. All patients received a 12-month Revlimid® maintenance. The median age of patients was 58 years (range: 30–65 years) and ISS stage distributions from stage 1 to 3 were 32%, 48.5%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and 19.5% respectively. To identify high-risk groups, standard fluorescence in situ hybridization (FISH) was performed on all patients. All patients, who achieved at least a very good partial response were also evaluated by sequencing-based Minimal residual disease (MRD) measurement. Details regarding the sample characteristics, FISH and MRD are provided in the supplementary data. All study participants provided written informed consent. As we were provided with all the high-risk markers, including MRD, ISS and FISH, we used our own dataset as the test dataset. We used the Multiple Myeloma Research Foundation's CoMMpass study, a prospective, longitudinal, observational study with clinical data, and captured the DNA and RNA level genomic data as the training dataset. All RNAseq files up to IA8 were downloaded from dbGap and 532 samples, that had survival data, as well as RNAseq newly diagnosed MM samples were used. Survival data from CoMMpass study IA9 release were used.

Paired-End RNA Sequencing and Analysis

The Methods section of the Supplementary Appendix contains a detailed description of RNA sequencing. Shortly after the RNA purification, the library preparation was completed by using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England BioLabs, Ipswich, MA), and was then converted into a DNA library following the manufacturer's protocol, with no modifications. The resulting libraries were then sequenced on the Illumina HiSeq 2000 (Illumina, San Diego, CA). The MMRF CoMMpass study samples were prepared using the Illumina TruSeq RNA library kit. Sequencing was performed on Illumina HiSeq2000 or HiSeq2500 instruments at TGen. Short reads obtained from the sequencing were mapped to the reference human genome GRCh38.

Copy Number Analysis

SNP array hybridization on the CytoScan HD array (Affymetrix) platform was performed according to the manufacturer protocols. 223 samples where we have RNAseq and clinical data were analyzed using Copy number analysis was performed using ChAS and Genotyping Console from Affymetrix and all downstream analysis was performed using R (v3.2.2). Dosage effect was calculated using iGC from Bioconductor.

Estimating the lincRNA Expression and Statistical Analysis

Transcript level counts were estimated by using the lightweight alignments tool, Salmon[20] and converted to gene-level estimates by summing the estimated transcript raw counts from Salmon within each gene. The efficiency of this method against traditional counting has been described in previous publications[21]. lincRNAs and protein coding genes were selected by using raw counts and genes with zero counts for all the samples were removed before normalization (See Supplementary Methods). 7277 lincRNAs and 19465 protein coding genes were used for downstream analysis. Library sizes were scale-normalized by the TMM method[22] prior to the voom[23] analysis. The Limma[24] pipeline was used for differential expression analysis. To evaluate the prognostic value of each lincRNA, we performed cox regression test in R (v 3.2.2) with survival (v 2.38-3), My.stepwise (v 0.1.0) and multtest (v 2.28.0) packages. The Kaplan-Meier plots showed the effects of risk scores, using the log-rank test. A multivariate cox proportional-hazard analysis for each risk score in a model, including MRD, ISS and cytogenetic risks were calculated on the independent test dataset to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

show independent effect of lincRNAs. Progression-free survival (defined as the time to a first relapse) and overall survival were used as endpoints. Detailed information about analysis steps is provided in the Supplementary Appendix.

Supervised lincRNA signature

We used relatively expressed (median expression > 1 after TMM+Voom normalization in the training CoMMpass dataset) lincRNAs (n=506) to create the lincRNA signature in MM. For each lincRNA, we used Cox proportional hazards regression model to identify significant (p value < 0.05) lincRNAs in the univariate analysis then we used to re-sampling-base multiple hypothesis testing with 1000 bootstrapping. The only lincRNAs that had p value < 0.05 were used to create multivariate model; 78 lincRNA from the first step were analyzed with step wise variable selection using the PFS data to create the final model. The risk score was calculated by the sum of the expressions multiplied by the sign of the coefficients from the multivariate model, if the high-expression group has poor outcome then the sign is set to positive, otherwise sign is set to negative. A total of 14 lincRNAs were included in the final model; low expressions of 7 (LINC00324, IQCH-AS1, TSIX, RP11-147L13.12, LINC00996, GLIDR, AC137934.1) as well as high expression of 7 (AP006621.5, FIRRE, RP11-1094M14.11, RP11-495P10.1, LINC01534, LINC01562, LINC01234)

RESULTS

LincRNAs are differentially expressed in MM and affected by copy number alterations

We compared expression of 7277 lincRNAs between purified MM cells from 308 patients and 16 purified normal plasma cells (NPC) from IFM/DFCI dataset and observed 869 differentially expressed lincRNAs in MM (Figure 1A, Supp. Figure 3, Supp. Figure 4 and Supp. Table 2). Of the differentially expressed lincRNAs, 45.5% (n= 395) were down regulated and 54.5% were (n= 474) up regulated in MM. There were 188 lincRNAs with at least 2 log2 fold difference between NPC and MM cells. Compared to protein coding genes, a greater proportion of the differentially expressed lincRNAs were up regulated in MM (42.7% vs 54.5%). Down / Up regulated gene ratio for lincRNAs were strongly effected by known copy number alterations in MM. The ratio was high for up regulated genes in hyperdiploid chromosomes 3,5,7,9 and 15 whereas chromosome 14 showed strong bias for down regulated lincRNAs (Figure 1B) suggesting a dosage effect dependencies for lincRNAs[25]. We further evaluated the dosage effects using copy number data, from Cytoscan HD SNP array, available for 226 patients. 153 lincRNAs showed significantly (FDR < 0.1) increased expression associated with copy number changes and 24 lincRNAs showed significantly (FDR < 0.1) decreased expression associated with copy number loss (Figure 1A, Supplementary Table 3 and 4).

Supervised prediction of outcome with lincRNAs

We next sought to develop a robust prognostic model using lincRNAs. Using the Compass data, a patient's risk score was calculated by summing 14 predictors, weighted by the signs of their marginal associations with the outcome (+ sign used for high expression high-risk lincRNAs and – sign used for low expression high-risk predictors). Patients were classified as a high-risk if their risk score was higher than 80th percentile. As seen in the Suppl. Figure

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

1, patients in low-risk groups had significantly superior PFS (HR: 3.67 [2.65–5.08], p value < 2e-16) and OS (HR: 3.80 [2.38– 6.07], p value = 1.5e-09) (Supplementary Figure 1) as compared to the high-risk group. To validate our signature, we utilized RNAseq data from 308 MM patient from independent IFM/DFCI clinical trial using the same pipeline and algorithms applied to the training dataset. We calculated the risk score for the test dataset using the 14 lincRNA signature. As seen in Figure 2A and 2B the lincRNA signature identified patients with high- and low-risk MM with significant difference in PFS (HR= 2.06 [1.44–2.96], p value = 5.7-e05) and OS (HR= 2.76 [1.51–5.05], p value = 0.00057).

lincRNAs identifies differential outcomes within the risk groups

To assess the added value of lincRNAs in predicting survival, we studied the impact of lincRNAs within known MM risk groups (ISS stage, high-risk cytogenetic and MRD status) in our test population. We applied the lincRNA risk score prediction to each of the good and poor risk groups defined by each of the 3 risk factors (Figure 3A–C for PFS and Figure 4A–C for OS) in the test dataset. We have recently shown that patients achieving MRD negative status, using NSG-based method that can detect 1 MM cell in 10^6 cells, have significantly superior PFS and OS. As seen in Figure 3A and 4A, within the subgroups of patients with both MRD negative and MRD positive status, lincRNA signature is further able to identify patients with good and poor risk myeloma. For example, the 4-year OS for MRD+ patients is 75%; within this group of patients, those with a good risk of lincRNA expression signature has 4 year PFS of 82% while those with poor risk lincRNA signature has 52% (p value = 0.0035). Similarly, patients with high vs low risk ISS and cytogenetic groups are each further separated into good and poor risk groups using lncRNA signature (Figures 3B, 4B and 3C,4C respectively). For example, patients with high-risk ISS, lincRNA-based good and poor risk groups have 45% and 13% 4 year PFS and 83% and 57% 4 year OS respectively (p value = 0.00012 and 0.0063); and in patients with low risk cytogenetics, lincRNA-based good and poor risk groups have 48% and 30% 4 year PFS and 90% and 64% 4 year OS respectively.

We next used multivariate Cox regression model to evaluate the known risk features with lincRNA prediction. In a multivariate analysis that included MRD (positive vs. negative), ISS (stage 1 vs. stage 2&3), cytogenetic risk (t (4;14), t(14;16), del17p vs. others), and lincRNA-based risk category (high vs. standard). Besides MRD measurement, lincRNA prediction showed a significant association with the EFS and OS (Figure 5 and Supp. Table 1).

Combining lincRNAs and protein coding genes can improve the prediction

To understand the effect of lincRNAs over protein coding genes, we utilized the EMC92 gene expression signature[26] in RNAseq platform for both our IFM/DFCI and MMRF CoMMpass dataset. EMC92 is the latest gene expression signature developed using the microarray based technology and, similar to other signatures, shows the highest prediction power in MM. For EMC-92; Probe-set IDs were converted into gene symbols and expression values converted into z scores, and original weights were used for each gene. As described in the original report, ~22% of patients were identified as high-risk. The EMC-92 signature could separate high-risk patients from low-risk patients for both OS and PFS using

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

both datasets (Supplementary Figure 2). We next combined lincRNAs and EMC-92 signature to show a combined genomics risk model for MM patients. As seen in Figure 6A and 6B, patients with both high-risk lincRNA and GEP signature have the lowest PFS and OS while those with high-risk in one group and low-risk in other have an intermediate outcome; while the patients with both low-risk lincRNA and low-risk GEP signatures have the best outcome, suggesting that lincRNA plays a role in driving clinical behavior of the disease.

DISCUSSION

MM is a heterogeneous disease with multiple subtypes and genomic alterations. The clinical outcome is determined by various risk predictors including, cytogenetic and FISH-identified abnormalities, as well as gene expression signatures. Recent reports have also described role for miRNA as well as mutations and copy number alteration in driving both biology and clinical course in myeloma. Here, we report lincRNA as an important genomic component that has significant impact on disease behavior. A differential expression of lincRNAs is observed between MM cells and normal plasma cells. Even though role of previously studied lincRNAs such as PVT1 (non-protein coding oncogene) and MIR22HG (micro RNA 22 host gene), LUCAT1, MEG3 has been reported, the precise biological role of majority of lincRNA remains unclear. This study now provides rationale for further detailed functional studies of lincRNAs and their impact on MM biology with an eye on potential therapeutic application in MM.

Here, we investigated the impact of lincRNAs on clinical outcome in MM using supervised approach, created using PFS, performed well on an independent validation data set with large sample size and longer follow-up. It is also important to note that the validation patient cohort had different therapeutic intervention than training patient cohort, suggesting that this lincRNA risk signature is not related with one specific intervention. The strength of the lincRNA-based risk stratification is also confirmed by significant separation of the 2 groups for OS even with only 16% of events with over 4-year median follow-up. Further significance of the lincRNA signature and its possible clinical utility is confirmed by its validation in 3 different risk groups. Its application to patients with ISS stage I and II/III provides further identification of patients with good and poor survival. Similarly, poor risk patients identified by FISH are further classified in 2 groups with significant survival difference.

The evaluation of minimal residual disease (MRD) is one of the most significant recent developments in clinical management of myeloma. A number of studies, including a large meta-analysis, have identified a significantly superior outcome in patients achieving MRD negative status[27]. Application of lincRNA signature to MRD negative patients identifies 2 groups with significantly different EFS and OS outcome. Thus, if a patient with good risk lincRNA achieves MRD negative status, he will be predicted to have an excellent survival outcome. In fact, a multivariate model with known risk features and lincRNA risk prediction showed that lincRNAs can be an independent molecular predictors and should be evaluated from clinical perspective.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Previously several gene expression signatures have been offered by various studies based on microarray platforms[26, 28–30]. Here we also combined lincRNAs with EMC-92 gene expression signature to understand their importance for potential clinical application. The combined lincRNA and protein coding genes prediction model performed better than both individual approaches. It is important to note that the EMC92 signature was developed using microarray-based technology and its performance using RNA-seq may not be optimized. However, performance of EMC-92 signature in our study here using RNAseq was acceptable in 2 different datasets. These results, besides providing a superior prediction model, also highlights that integrated analysis of various genomic correlates may have an added value for outcome prediction. It also highlights a possibility that functionally lincRNA may have role in cellular biology besides regulating the protein coding genes.

Although this study clearly provides evidence that lincRNAs are significant and need to be evaluated with other genomic features, further laboratory investigation is necessary to understand their biological interactions in the disease process and their roles as potential targets. In addition to which, due to their low-level abundance, compared to protein coding genes, it is also necessary to perform deep RNA sequencing to be able to study lincRNAs.

In conclusion, lincRNAs are dysregulated in MM and clearly provide valuable information on disease outcome. Incorporation of combined protein coding genes and lincRNA signature along with traditional risk features can identify patients with ultra-high-risk, as well as identify those who could potentially have excellent chances of survival.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

NIH grants P01-155258 and P50-100707, Department of Veterans Affairs Merit Review Award I01 BX001584-01

Research Support: NIH grants P01-155258 and P50-100707, Department of Veterans Affairs Merit Review Award I01 BX001584-01

References

1. Morris KV, Mattick JS. The rise of regulatory RNA. Nat Rev Genet. 2014 Jun; 15(6):423–437. [PubMed: 24776770]
2. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012 Sep; 22(9):1775–1789. [PubMed: 22955988]
3. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012 Sep; 22(9): 1760–1774. [PubMed: 22955987]
4. Huarte M. The emerging role of lncRNAs in cancer. Nat Med. 2015 Nov; 21(11):1253–1261. [PubMed: 26540387]
5. Zheng GX, Do BT, Webster DE, Khavari PA, Chang HY. Dicer-microRNA-Myc circuit promotes transcription of hundreds of long noncoding RNAs. Nat Struct Mol Biol. 2014 Jul; 21(7):585–590. [PubMed: 24929436]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

6. Zhang X, Gejman R, Mahta A, Zhong Y, Rice KA, Zhou Y, et al. Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. Cancer Res. 2010 Mar 15; 70(6):2350–2358. [PubMed: 20179190]
7. Fulciniti M, Amodio N, Bandi RL, Cagnetta A, Samur MK, Acharya C, et al. miR-23b/SP1/c-myc forms a feed-forward loop supporting multiple myeloma cell growth. Blood Cancer J. 2016 Jan 15.6:e380. [PubMed: 26771806]
8. Zarate R, Boni V, Bandres E, Garcia-Foncillas J. MiRNAs and LincRNAs: Could they be considered as biomarkers in colorectal cancer? Int J Mol Sci. 2012; 13(1):840–865. [PubMed: 22312290]
9. Yang N, Chen J, Zhang H, Wang X, Yao H, Peng Y, et al. LncRNA OIP5-AS1 loss-induced microRNA-410 accumulation regulates cell proliferation and apoptosis by targeting KLF10 via activating PTEN/PI3K/AKT pathway in multiple myeloma. Cell Death Dis. 2017 Aug 10.8(8):e2975. [PubMed: 28796257]
10. Gao D, Lv AE, Li HP, Han DH, Zhang YP. LncRNA MALAT-1 Elevates HMGB1 to Promote Autophagy Resulting in Inhibition of Tumor Cell Apoptosis in Multiple Myeloma. J Cell Biochem. 2017 Oct; 118(10):3341–3348. [PubMed: 28295550]
11. Ronchetti D, Manzoni M, Todoerti K, Neri A, Agnelli L. In Silico Characterization of miRNA and Long Non-Coding RNA Interplay in Multiple Myeloma. Genes (Basel). 2016 Nov 29.7(12)
12. Meng YB, He X, Huang YF, Wu QN, Zhou YC, Hao DJ. Long Noncoding RNA CRNDE Promotes Multiple Myeloma Cell Growth by Suppressing miR-451. Oncol Res. 2017 Aug 07; 25(7):1207–1214. [PubMed: 28276319]
13. Li B, Chen P, Qu J, Shi L, Zhuang W, Fu J, et al. Activation of LTBP3 gene by a long noncoding RNA (lncRNA) MALAT1 transcript in mesenchymal stem cells from multiple myeloma. J Biol Chem. 2014 Oct 17; 289(42):29365–29375. [PubMed: 25187517]
14. Ronchetti D, Agnelli L, Taiana E, Galletti S, Manzoni M, Todoerti K, et al. Distinct lncRNA transcriptional fingerprints characterize progressive stages of multiple myeloma. Oncotarget. 2016 Mar 22; 7(12):14814–14830. [PubMed: 26895470]
15. Cho SF, Chang YC, Chang CS, Lin SF, Liu YC, Hsiao HH, et al. MALAT1 long non-coding RNA is overexpressed in multiple myeloma and may serve as a marker to predict disease progression. BMC Cancer. 2014 Nov 04.14:809. [PubMed: 25369863]
16. Du Z, Sun T, Hacisuleyman E, Fei T, Wang X, Brown M, et al. Integrative analyses reveal a long noncoding RNA-mediated sponge regulatory network in prostate cancer. Nat Commun. 2016; 7:10982. [PubMed: 26975529]
17. Beck D, Thoms JAI, Palu C, Herold T, Shah A, Olivier J, et al. A four-gene LincRNA expression signature predicts risk in multiple cohorts of acute myeloid leukemia patients. Leukemia. 2017 Jul 04.
18. Mendell JT. Targeting a Long Noncoding RNA in Breast Cancer. N Engl J Med. 2016 Jun 09; 374(23):2287–2289. [PubMed: 27276568]
19. Attal M, Lauwers-Cances V, Hulin C, Leleu X, Caillot D, Escoffre M, et al. Lenalidomide, Bortezomib, and Dexamethasone with Transplantation for Myeloma. N Engl J Med. 2017 Apr 06; 376(14):1311–1320. [PubMed: 28379796]
20. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017 Apr; 14(4):417–419. [PubMed: 28263959]
21. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015; 4:1521. [PubMed: 26925227]
22. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010; 11(3):R25. [PubMed: 20196867]
23. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014; 15(2):R29. [PubMed: 24485249]
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20.43(7):e47. [PubMed: 25605792]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

25. Samur MK, Shah PK, Wang X, Minvielle S, Magrangeas F, Avet-Loiseau H, et al. The shaping and functional consequences of the dosage effect landscape in multiple myeloma. BMC Genomics. 2013 Oct 02.14:672. [PubMed: 24088394]

26. Kuiper R, Broyl A, de Knegt Y, van Vliet MH, van Beers EH, van der Holt B, et al. A gene expression signature for high-risk multiple myeloma. Leukemia. 2012 Nov; 26(11):2406–2413. [PubMed: 22722715]

27. Munshi NC, Avet-Loiseau H, Rawstron AC, Owen RG, Child JA, Thakurta A, et al. Association of Minimal Residual Disease With Superior Survival Outcomes in Patients With Multiple Myeloma: A Meta-analysis. JAMA Oncol. 2017 Jan 01; 3(1):28–35. [PubMed: 27632282]

28. Shaughnessy JD Jr, Zhan F, Burington BE, Huang Y, Colla S, Hanamura I, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. Blood. 2007 Mar 15; 109(6):2276–2284. [PubMed: 17105813]

29. Mulligan G, Mitsiades C, Bryant B, Zhan F, Chng WJ, Roels S, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood. 2007 Apr 15; 109(8):3177–3188. [PubMed: 17185464]

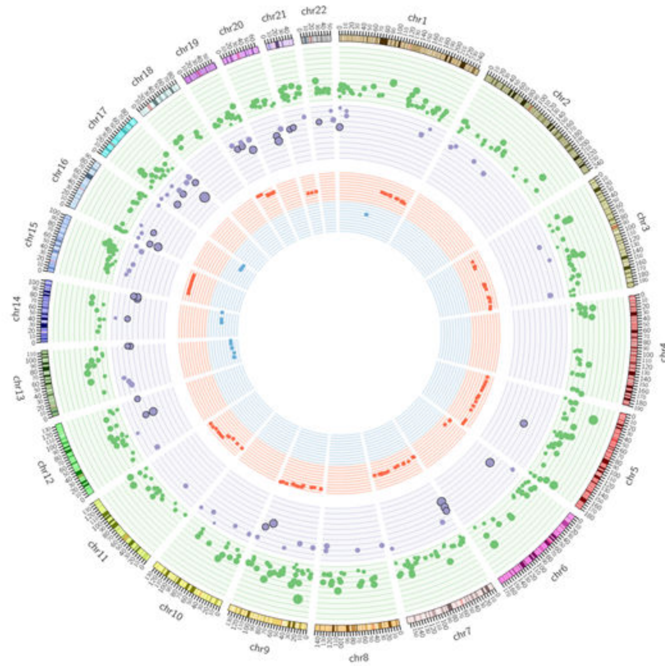
30. Decaux O, Lode L, Magrangeas F, Charbonnel C, Gouraud W, Jezequel P, et al. Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myelome. J Clin Oncol. 2008 Oct 10; 26(29): 4798–4805. [PubMed: 18591550]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



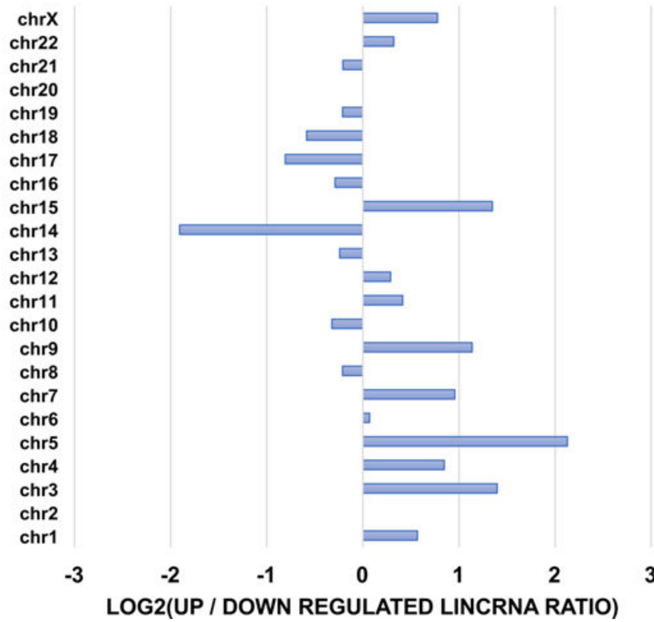


Figure 1. A genomic landscape of differentially expressed lincRNAs and their copy number in MM

a) A Circos plot shows differentially expressed lincRNAs across the genome. Green circle shows up-regulated lincRNAs and purple circle shows down regulated lincRNAs. y-axis (max=6, min=−6) and the glymph sizes were adjusted using log fold change values. Red circle and blue circles show the location of dosage effect dependent lincRNAs for gain and deletions respectively. y-axis (max=0.55) shows the proportion of MM samples that have deletion of gain for each lincRNAs. **b)** log2 ratio of up and down regulated lincRNAs in each chromosome.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

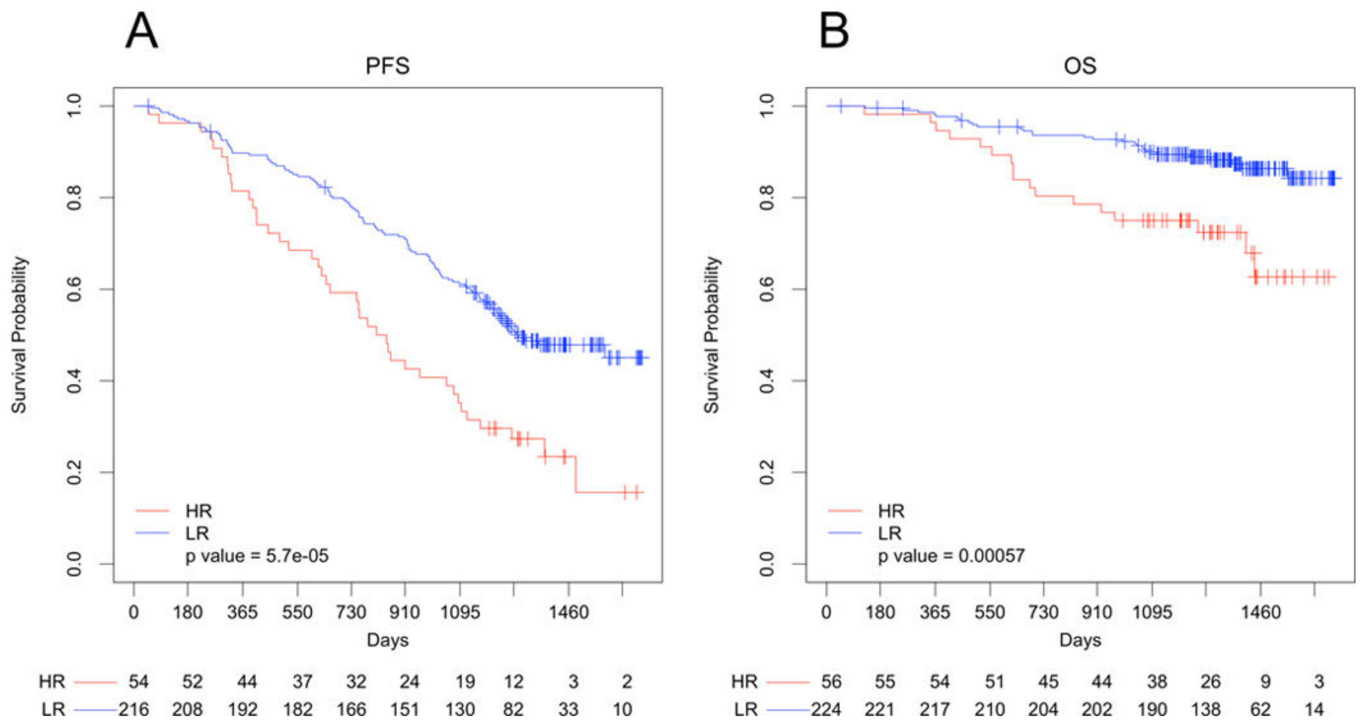


Figure 2. Outcome prediction on IFM/DFCI validation dataset

a) PFS probability of high and low-risk patients as classified by lincRNA risk model **b**) OS probability of high and low-risk patients as classified by lincRNA risk model

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

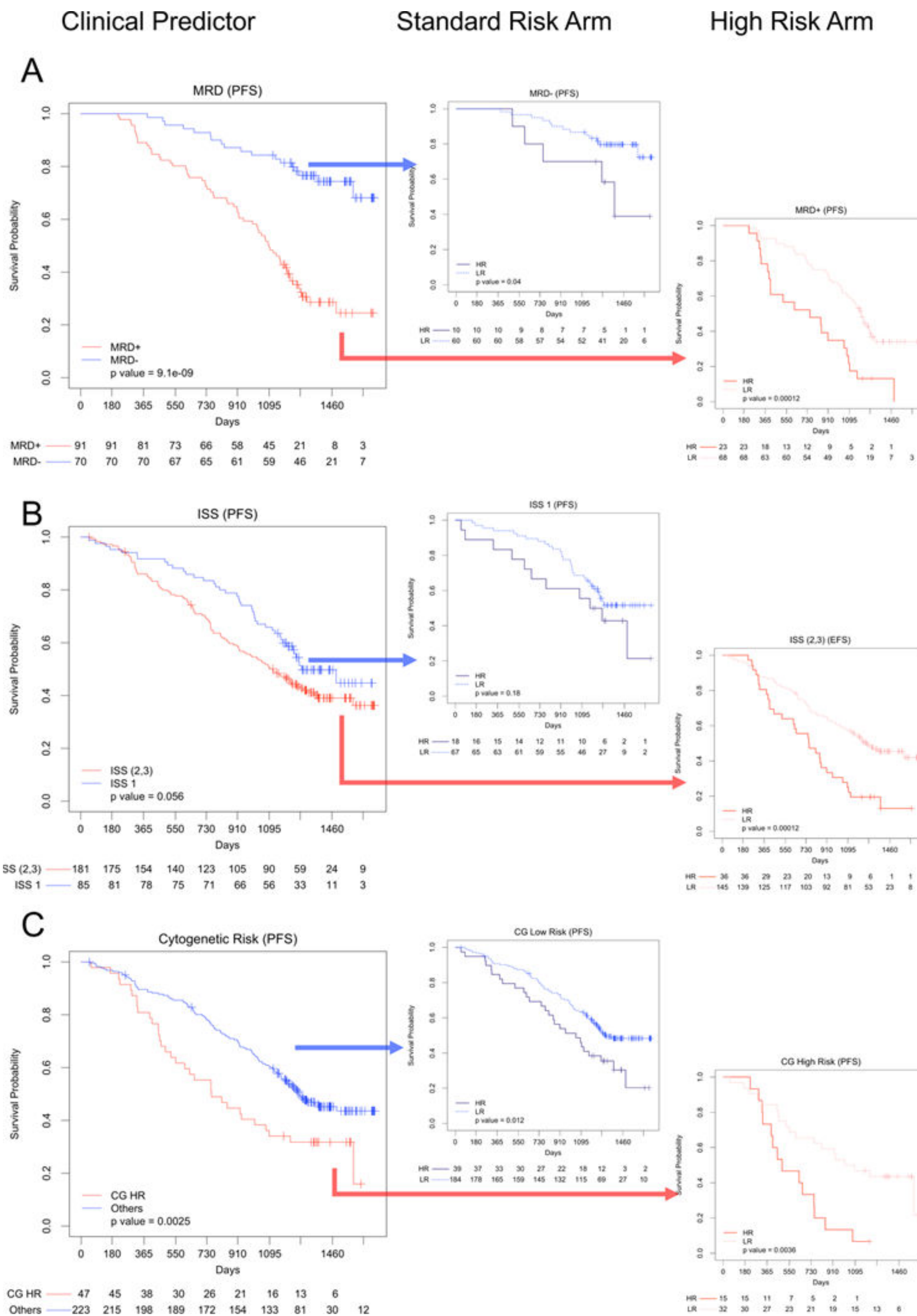


Figure 3. Added value of lincRNA signature on known clinical risk features for PFS

a) MRD alone, MRD- separated by lincRNA signature and MRD+ separated by lincRNA signature **b)** ISS alone, ISS 1 and ISS 2/3 separated by lincRNA signature **c)** Cytogenetic risk groups (del17p, t(4;14) and t(14;16) alone, Standard risk and high-risk separated by lincRNA signature respectively)

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

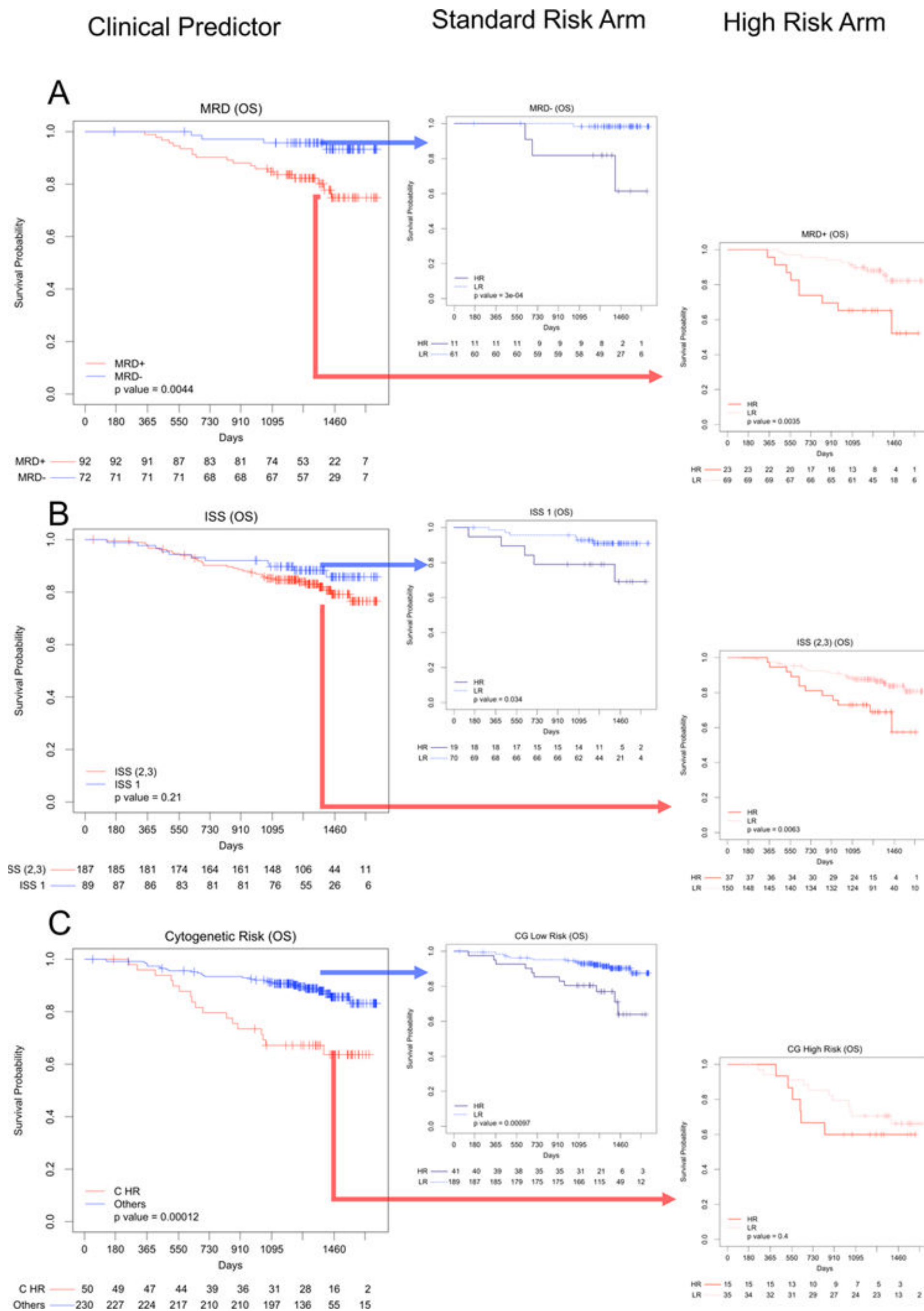


Figure 4. Added value of lincRNA signature on known clinical risk features for OS

a) MRD alone, MRD- separated by lincRNA signature and MRD+ separated by lincRNA signature **b)** ISS alone, ISS 1 and ISS 2/3 separated by lincRNA signature **c)** Cytogenetic risk groups (del17p, t(4;14) and t(14;16) alone, Standard risk and high-risk separated by lincRNA signature respectively)

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

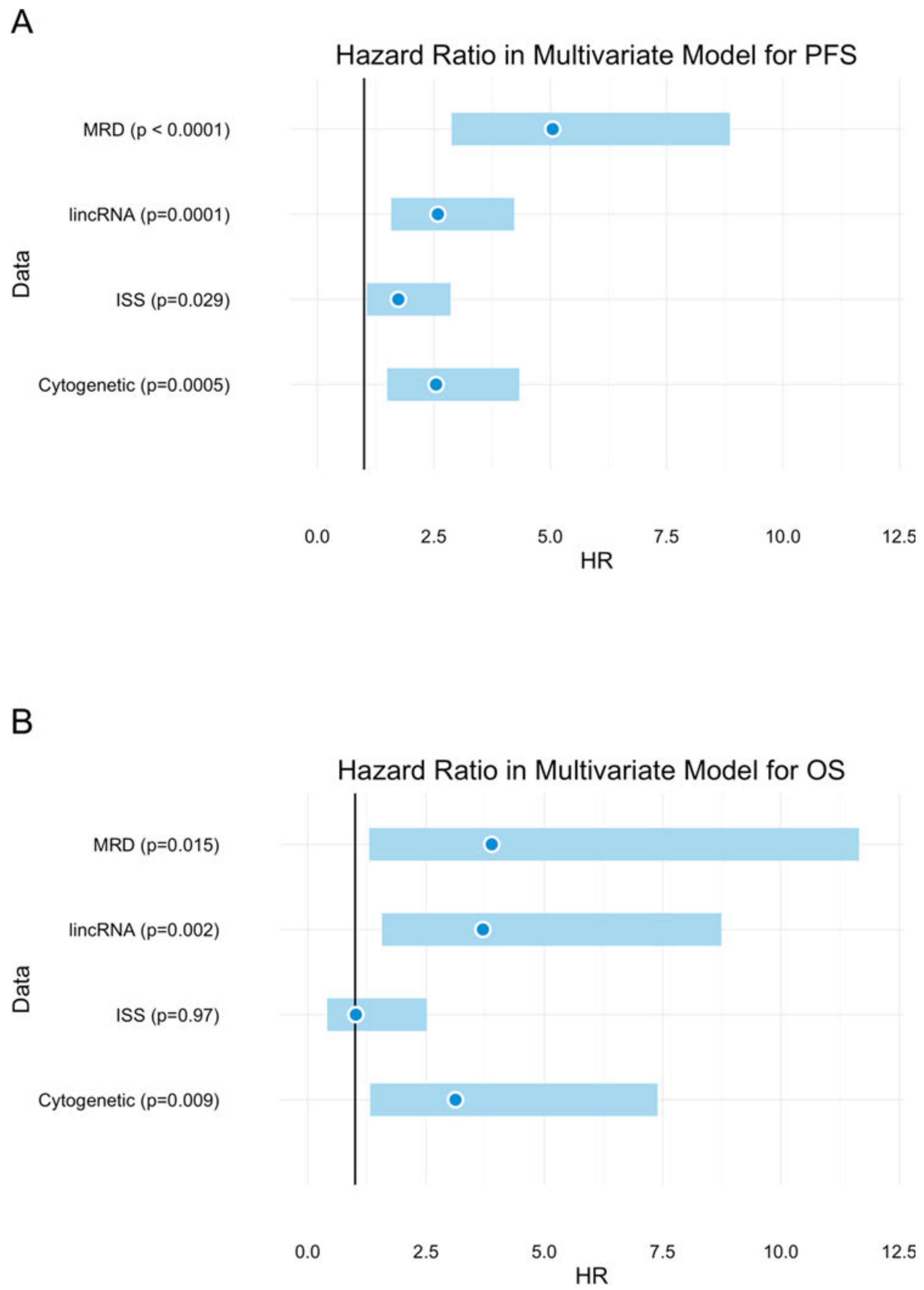


Figure 5. Hazard ratios of variables for the multivariate model in the IFM/DFCI dataset
a) PFS **b**) OS

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

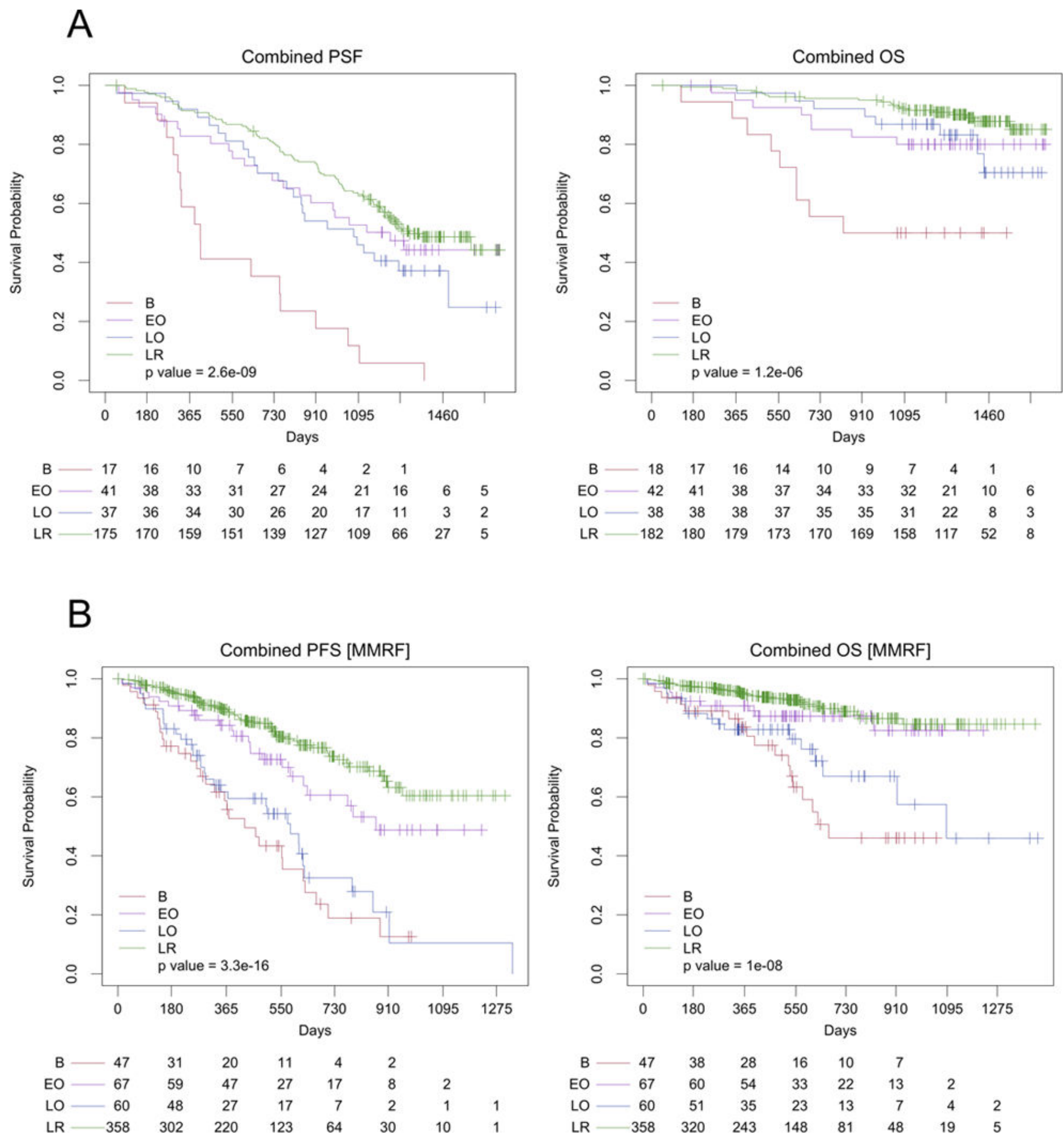


Figure 6. lincRNA risk prediction combined with EMC92 risk prediction

a) IFM/DFCI validation dataset and **b)** MMRF CoMMpass dataset for PFS and OS respectively. (B) high-risk for both predictor, (LO) lincRNA only high-risk, (EO) EMC92 only high-risk and (LR) low-risk for both predictors.