

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa Interunidades de Pós-Graduação em Bioinformática

Stellamaris Soares

**MODULAÇÃO DA EXPRESSÃO GÊNICA RENAL MEDIADA POR
CISPLATINA: uma visão integrativa entre genes e transcritos codificadores
e não codificadores de proteínas**

Belo Horizonte

2020

Stellamaris Soares

**MODULAÇÃO DA EXPRESSÃO GÊNICA RENAL MEDIADA POR CISPLATINA:
uma visão integrativa entre genes e transcritos codificadores e não codificadores de
proteínas**

Tese apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Doutora em Bioinformática.

Orientador: Dr. Carlos Alberto Tagliati

Coorientadora: Dra. Glória Regina Franco

Belo Horizonte

2020

043

Soares, Stellamaris.

Modulação da expressão gênica renal mediada por cisplatina: uma visão integrativa entre genes e transcritos codificadores e não codificadores de proteínas [manuscrito] / Stellamaris Soares. – 2020.

152 f. : il. ; 29,5 cm.

Orientador: Dr. Carlos Alberto Tagliati. Coorientadora: Dra. Glória Regina Franco.

Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática - Teses. 2. Néfrons - toxicidade. 3. Cisplatino. 4. RNA Longo não Codificante. 5. Transcriptoma. 6. Processamento Alternativo. I. Tagliati, Carlos Alberto. II. Franco, Glória Regina. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 573:004



UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa Interunidades de Pós-Graduação em Bioinformática da UFMG

ATA DE DEFESA DE TESE

STELLAMARIS SOARES

Às quatorze horas do dia **29 de maio de 2020**, reuniu-se, no Instituto de Ciências Biológicas da UFMG, a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho da doutoranda Stellamaris Soares, intitulado: "**Modulação da expressão gênica renal mediada por cisplatina: uma visão integrativa entre genes e transcritos codificadores e não codificadores de proteínas**", requisito para obtenção do grau de Doutora em **Bioinformática**.

Abrindo a sessão, o Presidente da Comissão, **Dr. Carlos Alberto Tagliati**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	Instituição	Indicação
Dr. Carlos Alberto Tagliati - Orientador	UFMG	Aprovada
Dra. Glória Regina Franco - Coorientadora	UFMG	Aprovada
Dra. Elaine Maria de Souza Fagundes	UFMG	Aprovada
Dra. Ana Maria Benko Iseppon	UFPE	Aprovada
Dr. Niels Olsen Saraiva Câmara	USP	Aprovada
Dra. Cristiane Paula Gomes Calixto	USP	Aprovada

Pelas indicações, a candidata foi considerada: **APROVADA**

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 29 de maio de 2020.



Documento assinado eletronicamente por **Stellamaris Soares, Usuário Externo**, em 02/06/2020,



às 22:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ana Maria Benko Iseppon, Usuário Externo**, em 02/06/2020, às 22:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Cristiane Paula Gomes Calixto, Usuário Externo**, em 03/06/2020, às 05:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Niels Olsen Saraiva Câmara, Usuário Externo**, em 03/06/2020, às 07:13, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Carlos Alberto Tagliati, Presidente de comissão**, em 03/06/2020, às 08:34, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Gloria Regina Franco, Membro de comissão**, em 03/06/2020, às 15:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Elaine Maria de Souza Fagundes, Professora do Magistério Superior**, em 04/06/2020, às 09:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0139932** e o código CRC **83A1507E**.

*Aos meus pais, Marcelo e Maria Augusta, e
aos meus irmãos, Leonardo e Anki.*

AGRADECIMENTOS

Aos meus pais e meus irmãos, tão essenciais em minha vida, agradeço por serem exatamente quem são e por fazerem parte da construção de quem eu sou. Agradeço por todo amor que existe entre nós, por todo companheirismo, por darem o melhor e mais bonito significado à palavra “família”. Agradeço por apoiarem meus sonhos sem restrições, por estarem sempre ao meu lado, pelos ensinamentos passados nas brincadeiras cheias de risadas, nas conversas bobas e nas sérias. Amo vocês!

Às minhas cunhadas, Carol e Durga, agradeço o amor, carinho e apoio, agradeço por serem presentes em minha vida e me ensinarem tanto. Agradeço por todo gesto e palavra de cuidado e de cura.

Aos meus sobrinhos, Bernardo e Breno, agradeço simplesmente por existirem. Ver a felicidade de vocês é uma das maiores alegrias que tenho. Agradeço por me mostrarem um amor tão grande, tão sincero e tão bonito.

Às minhas avós, Martha e Wanda – que no último ano foi brilhar em outros planos –, agradeço por serem exemplo de garra e por me ensinarem que um pouco de loucura não faz mal a ninguém e pode nos fazer mais felizes. À tia Maria – que também foi brilhar em outro plano neste ano que passou – e à tia Diva, agradeço por me permitirem retribuir o cuidado que sempre tiveram comigo.

Ao meu orientador, Carlos Tagliati, agradeço por ter aberto para mim as portas da pesquisa na UFMG. Agradeço por ter me apresentado à toxicologia de medicamentos, pela oportunidade de trabalhar nesta área e pelos ensinamentos ao longo de tantos anos.

À minha coorientadora, Glória Franco, agradeço por ter me recebido em seu grupo de pesquisa e por ter me orientado e apoiado em todas as fases do desenvolvimento deste projeto. Agradeço todo acolhimento e aconselhamento, agradeço por ser um exemplo de profissional que ama o que faz e que acredita que o compartilhamento do conhecimento é a melhor forma de crescimento para todos.

Aos professores do Laboratório de Genética Bioquímica (LGB), Andrea Macedo, Carlos Renato Machado e Erich Tahara, agradeço pelas discussões e contribuições para execução deste trabalho.

Aos colegas do Laboratório de Toxicologia, especialmente à Larissa Ribeiro e Marina Grossi que estiveram presentes desde o mestrado, agradeço por sempre auxiliarem e contribuírem com meus projetos e, principalmente, pela amizade.

Aos Gloriosos agradeço por terem me acolhido tão logo entrei para o grupo, por todo auxílio e contribuições para este trabalho, pelas muitas risadas que vocês me proporcionam, pela companhia dentro e fora do laboratório. Agradeço especialmente à Nayara Toledo por ter compartilhado seus conhecimentos e contribuído tanto para execução deste trabalho e, principalmente, pela amizade que construímos e levamos para além das portarias da UFMG. Agradeço pela companhia, pelos ouvidos sempre dispostos a ouvir, pelo carinho e por sempre ver o lado bom das pessoas.

Aos colegas do Bandeco, agradeço por todos os almoços regados de risadas e assuntos inapropriados para o momento, em especial à Ludmila Hufnagel, que se tornou amiga e companheira para todas as horas.

Aos colegas do LGB, agradeço pela companhia e apoio durante todos os anos de doutorado. Agradeço especialmente ao Bruno Resende por todos os momentos compartilhados, pelos desabafos e principalmente pelas alegrias.

Ao João Pedro, agradeço o companheirismo, a atenção, o apoio, a compreensão e o amor. Agradeço por permitir que eu faça parte da sua vida e por fazer parte da minha. A reta final do doutorado foi mais fácil e muito melhor com você.

Às minhas amigas da escola, Adriane Monteiro e Camila de Paula, e às da faculdade, Isabela Jardim, Gabriela Fonseca e Luísa Rodrigues, agradeço por todos esses anos de amizade. Em tantos anos, pudemos compartilhar muito de nossas vidas e espero que assim continue sendo.

Ao Curso de Verão em Bioinformática da UFMG, agradeço as oportunidades e todo aprendizado durante os 4 anos que participei da organização desse evento. Agradeço a atual Comissão Organizadora e aos antigos membros pelo ótimo trabalho em equipe que realizamos, fazendo com que a cada edição o Curso de Verão fosse ainda melhor que a anterior, agradeço também a amizade e compreensão de cada um de vocês.

Aos secretários da Pós-Graduação em Bioinformática, Sheila Santana e Tiago Silva, agradeço todo suporte e pela presteza em todos os momentos, com certeza, os melhores secretários que poderíamos ter.

Agradeço, por fim, o Programa Interunidades de Pós-Graduação em Bioinformática da UFMG, a Universidade Federal de Minas Gerais e a agência de fomento Capes.

*“Eu creio que a paixão é um ato de conhecimento
Quanto mais amo o que faço, mais invisto no que aprendo
Desejo que a invenção da profissão que movimento
Seja sempre consequência do que explode aqui dentro”*

Flaira Ferro

Apoio Financeiro

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES

RESUMO

A cisplatina é um quimioterápico largamente utilizado no tratamento de tumores sólidos e tem na nefrotoxicidade sua maior limitação de uso. Apesar dos esforços, os mecanismos de toxicidade ainda não estão completamente elucidados. Neste trabalho, a modulação da expressão gênica foi avaliada em rins de camundongos tratados com cisplatina a partir de dados publicamente disponíveis no *Gene Expression Omnibus* (GEO) sob os números de acesso GSE69652 e GSE106993. Resumidamente, os camundongos foram tratados com cisplatina ou com salina (controle). Após o tratamento os animais foram eutanasiados e tiveram os rins removidos, o RNA foi então extraído e sequenciado. A análise de expressão diferencial de transcritos identificou aumento na resposta celular ao dano de DNA, especialmente relacionada à apoptose por vias intrínsecas e diminuição da expressão de genes relacionados ao metabolismo basal celular, como metabolismo de aminoácidos e de ácidos graxos. A análise ao nível de transcritos revelou mudança nos padrões de *splicing* alternativo, onde genes codificadores expressaram transcritos em suas isoformas codificadoras e não codificadoras de proteínas. A expressão diferencial revelou, ainda, alteração em dezenas de RNAs longos não codificadores (lncRNA). A identificação destes lncRNAs adicionou novas informações aos mecanismos de regulação da morte celular por apoptose e da inflamação decorrentes da nefrotoxicidade induzida por cisplatina. Em conjunto, a expressão dos lncRNAs *lincRNA-p21* e *Gm26917* e de isoformas não codificadoras do gene codificador *Mdm2* revelaram forte regulação negativa de *Mdm2*, apesar da superexpressão desse gene, culminando na manutenção da apoptose. Pela primeira vez a transcrição *readthrough*, que acontece quando a transcrição continua após o sítio de terminação, está sendo relatada na nefrotoxicidade induzida por cisplatina. Os genes que sofreram transcrição *readthrough* estão envolvidos no processamento do mRNA. Evidências de que os RNAs de genes que sofrem transcrição *readthrough* não são traduzidos mostram que esse tipo de transcrição pode estar relacionado à modulação do *splicing* devido ao tratamento com cisplatina.

Palavras-chave: Nefrotoxicidade. Cisplatina. Análise de transcriptoma. *Splicing* alternativo. RNAs longos não codificadores. Transcrição *readthrough*.

ABSTRACT

Cisplatin is a chemotherapy drug widely used in the treatment of solid tumors and has in nephrotoxicity its greatest limitation. Despite efforts, the mechanisms of toxicity are not yet completely elucidated. In this work, the modulation of gene expression was evaluated in the kidneys of mice treated with cisplatin using publicly available data in the Gene Expression Omnibus (GEO) under access numbers GSE69652 and GSE106993. Briefly, the mice were treated with cisplatin or saline (control). At the end of the treatment, the animals were euthanized and the kidneys were removed, then the RNA was extracted and sequenced. Differential transcript expression analysis identified an increase in the cellular response to DNA damage, especially related to apoptosis by intrinsic pathways and a decrease in the expression of genes related to cellular basal metabolisms, such as amino acid and fatty acid metabolism. The analysis at the level of transcripts revealed a change in alternative splicing patterns, where coding genes expressed transcripts in their coding and non-coding protein isoforms. The differential expression also revealed changes in dozens of long non-coding RNAs (lncRNA). The identification of these lncRNAs added new information to the mechanisms of regulation of cell death by apoptosis and inflammation resulting from cisplatin-induced nephrotoxicity. Together, the expression of the *lincRNA-p21* and *Gm26917* lncRNAs and non-coding isoforms of the *Mdm2* protein-coding gene, revealed strong negative regulation of *Mdm2*, despite the overexpression of this gene, culminating in the maintenance of apoptosis. For the first time, readthrough transcription, which occurs when transcription continues after the termination site, is being reported in cisplatin-induced nephrotoxicity. The genes that have undergone readthrough transcription are involved in the processing of mRNA. Evidence that the RNAs of genes undergoing readthrough transcription are not translated shows that this type of transcription may be related to the modulation of splicing due to treatment with cisplatin.

Keywords: Nephrotoxicity. Cisplatin. Transcriptomic analysis. Long non-coding RNAs. Alternative splicing. Readthrough transcription.

LISTA DE ILUSTRAÇÕES

Figura 1. Morfologia e subestruturas renais.....	18
Figura 2. Modelo conceitual da lesão renal aguda.....	21
Figura 3. Estrutura química da cisplatina.....	23
Figura 4. Tipos de adutos formados entre cisplatina e DNA.....	25
Figura 5. Tecnologias para estudo do transcriptoma no decorrer dos anos.....	29
Figura 6. Tecnologias de sequenciamento de RNA.....	31
Figura 7. Exemplos de funções dos lncRNAs.....	35
Figura 8. Diferentes tipos de <i>splicing</i> alternativo.....	37
Figura 9. Principais etapas da terminação da transcrição e liberação de RNA polimerase II..	40
Figura 10. Fatores que podem afetar a eficiência de terminação da transcrição.....	41
Figura 11. Fluxo de trabalho resumido.....	44
Figura 12. Qualidade dos sequenciamentos.....	57
Figura 13. Perfil de distribuição de amostras.....	59
Figura 14. <i>Heatmaps</i> de transcritos diferencialmente expressos.....	61
Figura 15. Dispersão de <i>fold change</i> dos transcritos diferencialmente expressos.....	62
Figura 16. Transcritos diferencialmente expressos por categoria.....	63
Figura 17. Distribuição de genes por categorias de transcritos.....	64
Figura 18. Eventos de transcrição <i>readthrough</i>	67
Figura 19. <i>Heatmap</i> dos níveis de expressão dos DoGs.....	68
Figura 20. Processos biológicos enriquecidos dentre os genes codificadores de proteína superexpressos.....	70
Figura 21. Processos biológicos enriquecidos dentre os genes codificadores de proteína subexpressos.....	72
Figura 22. Processos biológicos enriquecidos dentre os genes que sofreram transcrição <i>readthrough</i>	73
Figura 23. Mecanismos de regulação de Mdm2.....	79

LISTA DE TABELAS

Tabela 1. Características do desenho experimental dos estudos analisados.....	55
Tabela 2. Identificação das amostras.....	56
Tabela 3. Percentual de mapeamento de <i>reads</i>	58

LISTA DE ABREVIATURAS E SIGLAS

AA	Ácido araquidônico
AIF	<i>Apoptosis-inducing factor</i>
ATP	Adenosina trifosfato
ATR	<i>Ataxia telangiectasia and Rad3-related</i>
CAGE	<i>Cap Analysis of Gene Expression</i>
cDNA	DNA complementar
CDS	<i>Coding sequence</i>
CP	Codificadores de proteínas
CPSF	<i>Cleavage and polyadenylation specificity factor</i>
CstF	<i>Cleavage stimulation factor</i>
Ctrl	<i>Cooper transporter protein 1</i>
DNA	Ácido desoxirribonucleico
DoG	<i>Downstream of a gene containing transcript</i>
EET	Ácidos epoxieicosatrienoicos
ENaC	<i>Epithelial sodium channel</i>
EST	<i>Expressed Sequence Tag</i>
FDA	<i>Food and Drug Administration</i>
GEO	<i>Gene Expression Omnibus</i>
GO	<i>Gene Ontology</i>
HSV-1	Vírus Herpes Simples
ICAM-1	<i>Intercellular adhesion molecule 1</i>
IGFBP7	<i>Insulin like growth factor binding protein 7</i>
IL	Interleucina
KIM-1	<i>Kidney injury molecule 1</i>
L-FABP	<i>Liver-type fatty acid binding protein</i>
lincRNA	<i>Long interspersed ncRNA</i>
lncRNA	RNA longo não codificador de proteína
LRA	<i>Lesão renal aguda</i>
MATE1	<i>Multidrug and toxin extrusion 1</i>
MCP-1	<i>Macrophage cationic peptide 1</i>
MIP-2	<i>Macrophage inflammatory protein 2</i>
miRNA	microRNA
mRNA	RNA mensageiro

NCBI	<i>National Center for Biotechnology Information</i>
ncRNA	RNA não codificador de proteína
NCP	Não codificadores de proteínas
NMD	<i>Nonsense mediated decay</i>
NEAT1	<i>Nuclear enriched abundant transcript 1</i>
NF-κB	<i>Nuclear factor kappa B subunit 1</i>
NGAL	<i>Neutrophil gelatinase-associated lipocalin</i>
OCT	<i>Organic cation transporter</i>
PAP	Poli-A-polimerase
pb	Pares de bases
PCA	Análise de componentes principais
pH	Potencial hidrogeniônico
PIDD	<i>p53-induced protein with death domain</i>
Pol II	RNA polimerase II
PRINS	<i>Psoriasis susceptibility-related RNA gene induced by stress</i>
PUMA-α	<i>p53-upregulated modulator of apoptosis</i>
PVT1	<i>Plasmocytoma variant translocation 1</i>
RANTES	<i>Regulated on activation, normal T cell expressed and secreted</i>
RNA	Ácido ribonucleico
RNA-Seq	Sequenciamento de RNA
RPKM	<i>Reads por quilobase por milhão de reads mapeadas</i>
SAGE	<i>Serial Analysis of Gene Expression</i>
SRA	<i>Sequence Read Archive</i>
TDE	Transcritos diferencialmente expressos
TFF3	<i>Trefoil factor 3</i>
TFG	Taxa de filtração glomerular
TGF-β	<i>Transforming growth factor beta</i>
TIMP2	<i>Tissue inhibitor of metalloproteinases 2</i>
TNFα	<i>Tumor necrosis factor alpha</i>
TPCP	Transcritos processados originados de genes codificadores de proteínas
TRV	Teste de razão de verossimilhança
UCSC	<i>University of California, Santa Cruz</i>

SUMÁRIO

1 INTRODUÇÃO.....	17
1.1 Os rins.....	17
1.2 Lesão renal aguda.....	20
1.3 Cisplatina.....	22
1.3.1 Nefrotoxicidade induzida por cisplatina.....	23
1.4 Transcriptômica e tecnologias associadas.....	27
1.5 Toxicogenômica.....	32
1.6 Genes não codificadores de proteínas.....	33
1.7 <i>Splicing</i> alternativo como mecanismo de variabilidade de transcritos.....	36
1.8 Transcritos originados de genes codificadores de proteínas.....	38
1.9 Transcrição <i>readthrough</i> – transcrição de término estendido.....	39
2 OBJETIVOS.....	43
2.1 Objetivo geral.....	43
2.2 Objetivos específicos.....	43
3 METODOLOGIA.....	44
3.1 Fluxo de trabalho resumido.....	44
3.2 Aquisição de dados de sequenciamento de RNA e controle de qualidade.....	44
3.3 Alinhamento de <i>reads</i> contra o genoma de referência.....	46
3.4 Pseudoalinhamento contra o transcriptoma de referência e estimativa de abundância de transcritos.....	46
3.5 Expressão diferencial de transcritos.....	47
3.6 Identificação de transcrição <i>readthrough</i>	48
3.7 Análise funcional.....	50
3.7.1 Genes codificadores de proteínas.....	50
3.7.2 Genes não codificadores de proteínas.....	52
3.7.3 Genes que sofreram transcrição <i>readthrough</i>	53
3.8 Visualização dos resultados.....	53
3.9 Recursos computacionais.....	53
4 RESULTADOS.....	54
4.1 Caracterização dos estudos analisados.....	54
4.2 Caracterização das amostras.....	57
4.3 Estudo do transcriptoma renal após tratamento com cisplatina.....	60
4.3.1 Perfil de expressão diferencial de transcritos.....	60
4.3.2 Transcrição <i>readthrough</i>	66

4.3.3 Análise funcional.....	68
4.3.3.1 Transcritos superexpressos.....	68
4.3.3.2 Transcritos subexpressos.....	71
4.3.3.3 Genes que sofreram transcrição <i>readthrough</i>	72
5 DISCUSSÃO.....	75
6 CONCLUSÕES.....	90
7 PERSPECTIVAS.....	91
REFERÊNCIAS.....	92
APÊNDICE A – Lista de fármacos nefrotóxicos.....	107
APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2.....	108
APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2.....	117
APÊNDICE D – Processos biológicos enriquecidos nos transcritos codificadores de proteínas superexpressos.....	131
APÊNDICE E – Processos biológicos enriquecidos nos transcritos processados oriundos de genes codificadores de proteínas superexpressos.....	132
APÊNDICE F – Processos biológicos enriquecidos nos transcritos codificadores de proteínas subexpressos.....	133
APÊNDICE G – Processos biológicos enriquecidos nos transcritos processados oriundos de genes codificadores de proteínas superexpressos.....	134
APÊNDICE H – Processos biológicos enriquecidos nos genes que sofreram transcrição <i>readthrough</i> – DoGs superexpressos.....	135
APÊNDICE I – Processos biológicos enriquecidos nos genes que sofreram transcrição <i>readthrough</i> – DoGs subexpressos.....	136
APÊNDICE J – Publicação relacionada à tese.....	137

1 INTRODUÇÃO

1.1 Os rins

Os rins são órgãos constituintes do sistema urinário, possuem formato assemelhado ao grão de feijão e localizam-se na cavidade retroperitoneal do corpo. Na borda côncava, encontra-se o hilo, por onde passam o suprimento nervoso, a artéria e veia renais, os vasos linfáticos e o ureter. Internamente, através de corte longitudinal, é possível identificar duas regiões principais, uma mais externa denominada córtex e outra mais interna denominada medula (Figura 1A) (HALL, 2017; CONSTANZO, 2018).

Os rins são responsáveis pela filtração do sangue, eliminando substâncias tóxicas ou desnecessárias ao organismo através da urina. Dentre essas substâncias estão produtos do metabolismo como a ureia, creatinina, ácido úrico e aquelas ingeridas, como medicamentos, pesticidas e aditivos alimentícios. Outra função fundamental dos rins é regular o equilíbrio de água e de eletrólitos, a fim de manter a homeostasia de todo o organismo. Para isso, a eliminação de água e eletrólitos é baseada no nível de ganho dos mesmos. A regulação do equilíbrio de água e eletrólitos faz parte também da regulação da pressão arterial, juntamente à secreção renal de hormônios e substâncias vasoativas. Além disso, os rins possuem outras funções essenciais ao bom funcionamento do organismo, tais como a regulação do equilíbrio ácido-base através da excreção de ácidos, secreção, metabolismo e excreção de hormônios e gliconeogênese em situações de jejum prolongado (HALL, 2017).

Para exercer todas suas funções, os rins recebem cerca de 25% do débito cardíaco, o maior valor dentre todos os órgãos. Os rins são órgãos altamente vascularizados, sendo o fluxo sanguíneo renal médio de 1.100 mL/min (AIRES, 2018). A unidade funcional dos rins é denominada néfron e cada rim humano possui entre 800.000 e 1.200.000 néfrons. Os néfrons são constituídos pelo corpúsculo renal em sua porção inicial e uma porção tubular, esta última é subdividida em túbulo proximal, alça de Henle, túbulo distal e ducto coletor. Cada uma das regiões dos néfrons possui funções específicas que, em conjunto, contribuem para a formação da urina (Figura 1B) (AIRES, 2018). A seguir, serão apresentadas, resumidamente, as etapas de formação da urina em cada uma das regiões do néfron.

A primeira etapa da formação da urina é a filtração glomerular. O corpúsculo renal é composto pelo glomérulo, uma rede de capilares, que é envolvido pela cápsula de Bowman. Ao passar pelo glomérulo, o plasma sanguíneo é filtrado para o espaço de Bowman. Esse

primeiro filtrado é denominado ultrafiltrado e tem composição similar à do sangue, no entanto, não contém células nem proteínas. A modificação do ultrafiltrado em urina envolve processos de reabsorção e secreção, que ocorrem na região tubular do néfron. No processo de reabsorção, substâncias que não devem ser excretadas são levadas do lúmen dos túbulos de volta à corrente sanguínea. Enquanto, no processo de secreção, as células tubulares retiram as substâncias que devem ser excretadas do sangue capilar peritubular e as adicionam ao filtrado no lúmen dos túbulos (CONSTANZO, 2018).

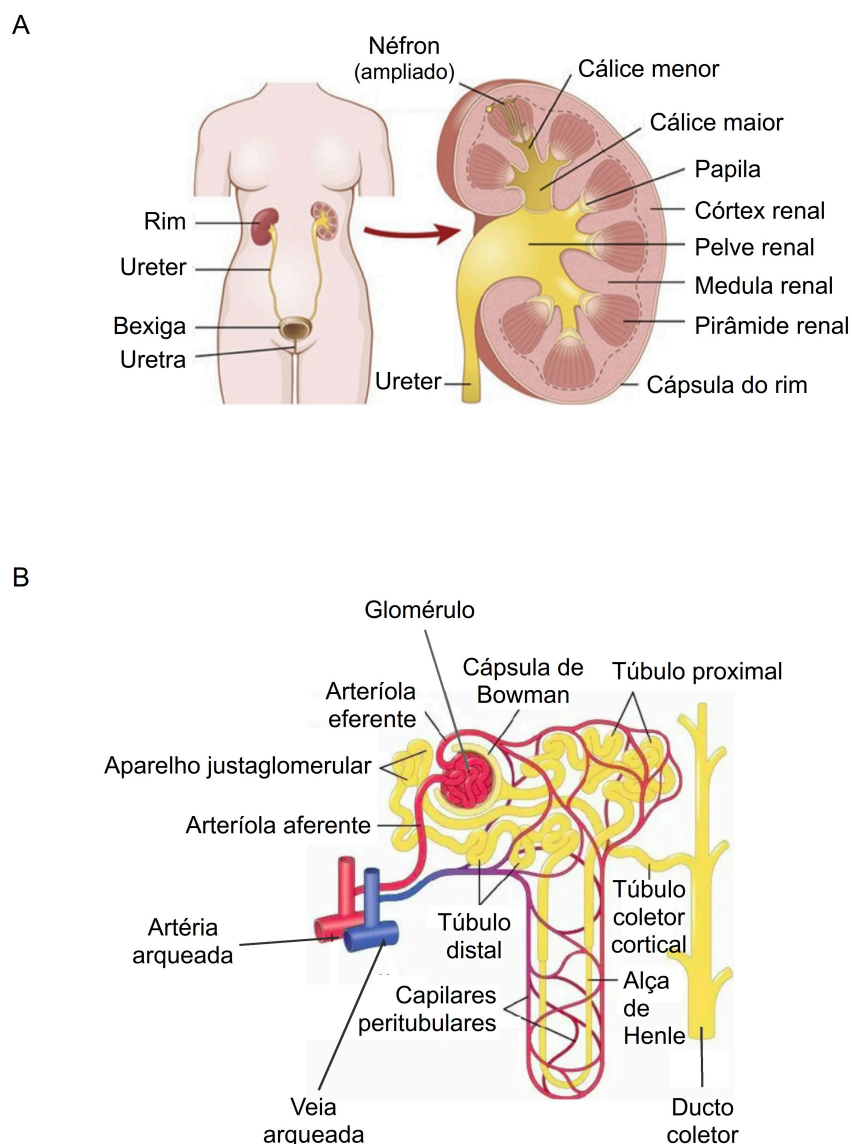


Figura 1. Morfologia e subestruturas renais. (A) Esquema do sistema urinário e regiões internas do rim. (B) Esquema de um néfron e sua microcirculação.

A porção que se segue ao glomérulo é o túbulo proximal, sendo os dois primeiros terços compostos por túbulos contorcidos e o último terço composto por túbulo reto. O túbulo contorcido proximal é responsável pela reabsorção de cerca de 80% do líquido do filtrado e pela reabsorção completa de glicose e aminoácidos. Cerca de 70% do cloreto e 80% dos íons de sódio são reabsorvidos no túbulo proximal, outros íons como potássio, magnésio, bicarbonato, cálcio e ureia e ácido úrico são reabsorvidos em quantidades variáveis. As células do túbulo proximal também realizam secreção de substâncias que devem ser rapidamente retiradas do organismo como, por exemplo, ácidos e bases orgânicos, sais biliares, íons hidrogênio e substâncias potencialmente danosas ao organismo, como medicamentos e toxinas (HALL, 2017; AIRES, 2018).

Na sequência, o filtrado segue pela alça de Henle, a qual possui três porções distintas, sendo elas o ramo fino descendente, ramo fino ascendente e ramo grosso ascendente. O ramo descendente fino é altamente permeável à água e moderadamente permeável a solutos como sais e ureia. Já os ramos ascendentes fino e grosso possuem baixa permeabilidade à água e alta taxa de reabsorção de sais, sendo que o ramo ascendente fino tem menor capacidade de reabsorção de solutos que o ramo ascendente grosso. Uma vez que o ramo ascendente grosso é praticamente impermeável à água e reabsorve ativamente grandes quantidades de soluto, o líquido tubular torna-se cada vez mais diluído à medida que segue em direção ao túbulo distal, o que permite que os rins concentrem ou diluam a urina de acordo com as necessidades do organismo (HALL, 2017).

A porção inicial do túbulo distal é composta por células epiteliais especializadas, denominada mácula densa, em contato íntimo com as células justaglomerulares que recobrem um segmento das arteríolas aferente e eferente, responsáveis pela entrada e saída do fluxo sanguíneo no glomérulo. Em conjunto, as células da mácula densa e as células justaglomerulares formam o complexo justaglomerular. Este complexo é responsável por um sistema de *feedback* que regula o fluxo sanguíneo renal e a taxa de filtração glomerular e, com isso, mantém os níveis de cloreto de sódio que chegam ao túbulo distal relativamente constante. A mácula densa detecta o aumento da entrada de água e cloreto de sódio originadas do aumento do fluxo sanguíneo renal e da taxa de filtração glomerular. Em resposta a esse aumento, a mácula densa secreta uma substância vasoativa levando a uma constrição da arteríola aferente que, por sua vez, reduz o fluxo sanguíneo renal fazendo com que a taxa de filtração glomerular volte ao normal (HALL, 2017; CONSTANZO, 2018).

O túbulo contorcido distal é rico em receptores de aldosterona e, em resposta a este hormônio, reabsorve sódio e secreta potássio e hidrogênio. Essa porção do néfron também

reabsorve bicarbonato, cálcio e potássio e secreta amônia. A porção inicial do túbulo contorcido distal é relativamente impermeável à água, já a porção final responde ao hormônio antidiurético, reabsorvendo pequenas quantidades de água, além de ser pouco permeável à ureia. Essas características fazem com que a ureia seja concentrada no líquido intratubular ao passar pelo túbulo contorcido distal (AIRES, 2018).

A porção final do túbulo distal e o ducto coletor possuem características funcionais similares, sendo ducto coletor o local final para o processamento da urina (HALL, 2017). O ducto coletor é impermeável à água, tornando-se permeável somente na presença do hormônio antidiurético. Essa região reabsorve íons cloreto, potássio, hidrogênio e bicarbonato e secreta amônia e bicarbonato. Em resposta a aldosterona, reabsorve íon sódio e secreta os íons potássio e hidrogênio (AIRES, 2018). O bom funcionamento renal é essencial ao funcionamento do organismo como um todo. Sendo assim, danos renais podem levar a sérias consequências ao indivíduo.

1.2 Lesão renal aguda

Os rins são alvos importantes de processos de toxicidade devido ao alto fluxo sanguíneo que recebem, fazendo com que sejam expostos regularmente a substâncias potencialmente danosas tais como medicamentos e seus metabólitos (PAZHAYATTIL; SHIRALI, 2014). Esses e outros fatores podem levar ao quadro de lesão renal aguda (LRA), a qual é caracterizada pela perda abrupta da função renal que pode ocorrer em poucos dias ou mesmo em algumas horas. O diagnóstico de LRA é baseado em alterações nos valores de creatinina sérica e ou volume de urina. Assim, a LRA é diagnosticada pelo aumento na creatinina sérica em nível maior ou igual a 0,3 mg/dL dentro de 48 h ou aumento maior ou igual a 1,5 vezes o valor de base, conhecido ou presumido, dos sete dias anteriores ou volume de urina menor que 0,5 mL/kg/h por 6 h (KDIGO ACUTE KIDNEY INJURY WORK GROUP, 2012).

A LRA pode ser adquirida fora e dentro do ambiente hospitalar. Quando adquirida fora do ambiente hospitalar, pode ser causada por infecções, toxinas e desidratação, geralmente associadas a doenças agudas. Já no ambiente hospitalar, respostas tóxicas a medicamentos e lesões renais isquêmicas devido à cirurgia, sepse ou trauma são as causas de muitos casos de LRA (ZUK; BONVENTRE, 2016). Medicamentos prescritos e outros medicamentos tradicionais (naturais) são a principal causa da LRA, resultando em quase 25% dos casos

(KELLUM; RONCO; BELLOMO, 2016). A nefrotoxicidade induzida por medicamentos é uma das principais causas de mortalidade e morbidade por LRA (NASH; HAFEEZ; HOU, 2002).

A progressão do quadro de LRA pode evoluir para doença renal crônica e insuficiência renal, podendo até mesmo levar à morte. Além disso, a LRA pode ser um fator de risco em outras doenças como, por exemplo, acidente vascular cerebral, sepse e doenças cardiovasculares (Figura 2) (LEVIN et al., 2017; FORTRIE; DE GEUS; BETJES, 2019). Ainda não estão completamente esclarecidos os motivos pelos quais pequenas alterações nos níveis de creatinina sérica levam ao aumento de complicações e da mortalidade hospitalar. Sobrecarga de volume, retenção de toxinas urêmicas, acidose, distúrbios eletrolíticos, aumento do risco de infecção e anemia são efeitos da diminuição da função renal e compreendem possíveis explicações para essa questão (KDIGO ACUTE KIDNEY INJURY WORK GROUP, 2012).

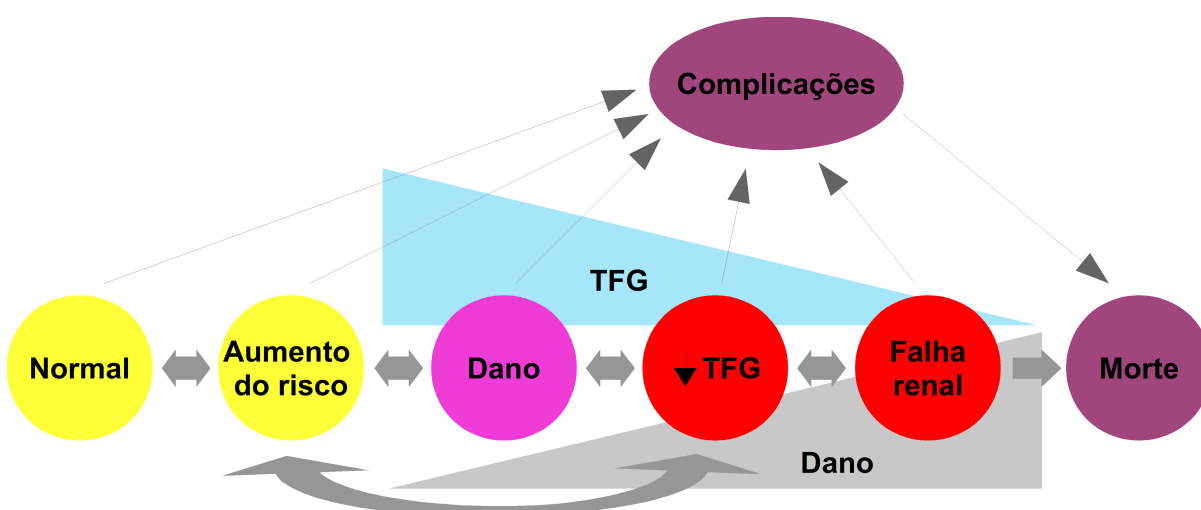


Figura 2. Modelo conceitual da lesão renal aguda. Círculos amarelos, antecedentes potenciais de LRA. Círculo rosa, estágio intermediário. Círculos vermelhos, estágios da LRA. Círculos lilás, resultados da LRA. Setas grossas cinzas entre os círculos, fatores de risco associados ao início da progressão da doença que podem ser afetadas ou detectadas por intervenções. LRA, lesão renal aguda. TGF, taxa de filtração glomerular.

Adaptado de Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney inter.*, Suppl. v. 2, p. 1-138, 2012.

Os atuais marcadores para diagnóstico de LRA, volume urinário e creatinina sérica, apresentam limitações relacionados à especificidade e sensibilidade. Na presença de dano, a concentração de creatinina sérica só é aumentada após a perda de aproximadamente 50% da função renal. Além disso, os níveis de creatinina sérica são dependentes de diversos fatores,

tais como alimentação, massa muscular, idade e uso de medicamentos (LASSNIGG et al., 2004; WEISS et al., 2019; YONEYAMA et al., 2019).

Em vista disso, novos biomarcadores têm sido constantemente investigados. O órgão regulador dos Estados Unidos da América, *Food and Drug Administration* (FDA), aprovou sete biomarcadores para uso em animais e, caso a caso, em humanos, sendo eles KIM-1 (*kidney injury molecule 1*), β 2-microglobulina, cistatina C, clusterina, TFF3 (*trefoil factor 3*), albumina e proteína total (ZUK; BONVENTRE, 2016). Além desses, outros biomarcadores continuam sob investigação para diagnóstico, progressão e severidade da LRA como, por exemplo, TIMP-2 (*tissue inhibitor of metalloproteinases 2*), IGFBP7 (*insulin like growth factor binding protein 7*) (MEERSCH et al., 2014; CHO et al., 2019), NGAL (*neutrophil gelatinase-associated lipocalin*), L-FABP (*liver-type fatty acid binding protein*) (YONEYAMA et al., 2019), interleucina (IL)-24 (TABATA et al., 2019).

A detecção precoce de LRA, ou seja, em um momento onde o nível de dano renal ainda não causa comprometimento funcional, é importante para prevenir o agravamento da lesão renal, pois permite o início precoce de intervenções nefroprotetoras (WEISS et al., 2019). Biomarcadores de lesão renal em nível celular podem ajudar a identificar pacientes que estejam em estágios mais iniciais de LRA, além de identificar maior número de pacientes com LRA (KDIGO ACUTE KIDNEY INJURY WORK GROUP, 2012). Assim, compreender os mecanismos moleculares envolvidos na LRA é fundamental para a identificação de biomarcadores precoces, bem como para identificar novos alvos terapêuticos, possibilitando o tratamento adequado e diminuindo o número de pacientes que evoluem para maiores problemas de saúde ou para morte.

1.3 Cisplatina

A cisplatina ($\text{Pt}(\text{NH}_3)_2\text{Cl}_2$) é um complexo inorgânico formado por um átomo central de platina (Pt), onde dois átomos de cloro (Cl) e duas moléculas de amônia ($-\text{NH}_3$) se ligam em orientação cis (Figura 3). Apesar de ser comum e erroneamente descrita como agente alquilante devido a seu mecanismo de ação, a cisplatina não possui nenhum grupo alquila (C_nH_{n+1}) em sua estrutura. Sintetizada pela primeira vez em 1845, a cisplatina foi inicialmente chamada de cloreto de Peyrone, tendo sua estrutura química elucidada em 1893 por Alfred Werner (ROSENBERG, 1980; DASARI; TCHOUNWOU, 2014).

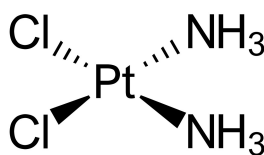


Figura 3. Estrutura química da cisplatina.

A primeira evidência de que a cisplatina poderia possuir características de interesse médico foi vista em 1965, quando Rosenberg e colaboradores conduziam experimentos com o objetivo de identificar se campos elétricos afetariam o crescimento de bactérias. Naquele momento, acreditava-se que platina era um composto inerte e, portanto, foram utilizados eletrodos constituídos por este metal. No entanto, observou-se a presença de compostos de platina no meio de cultura eletrolisado, observando-se que estes eram os causadores da inibição de divisão celular observada em *Escherichia coli* (ROSENBERG; VAN CAMP; KRIGAS, 1965). A partir desta descoberta, estudos sobre atividade antitumoral foram conduzidos utilizando diversos compostos de platina, incluindo a cisplatina. A regressão de tumor foi observada em camundongos transplantados com sarcoma, onde os autores sugeriram que compostos de platina seriam uma nova classe de agentes antitumorais (ROSENBERG et al., 1969).

Rapidamente, a cisplatina entrou em testes clínicos, ou seja, em humanos, sendo os primeiros pacientes tratados com cisplatina em 1971. Já em 1978, o uso de cisplatina para o tratamento de câncer de próstata foi aprovado pelo FDA (HIGBY et al., 1974; KELLAND, 2007). Atualmente, a cisplatina é utilizada em regimes de tratamento de diversos tumores sólidos como, por exemplo, cânceres de ovário, mama, testículo, próstata, bexiga, cabeça e pescoço, pulmão, esôfago, estômago, em melanoma, sarcomas, neuroblastoma e também em linfomas (MANOHAR; LEUNG, 2018). Apesar de toda sua utilidade terapêutica, a cisplatina pode induzir efeitos tóxicos significativos, limitando seu uso em pacientes susceptíveis.

1.3.1 Nefrotoxicidade induzida por cisplatina

A nefrotoxicidade foi uma característica observada já nos primeiros anos de estudo da cisplatina. O dano renal foi considerado um limitante de dose durante os estudos clínicos de fase I (HIGBY et al., 1974). A LRA é um dos maiores problemas relacionados ao tratamento com cisplatina, aparecendo entre 20% e 30% dos pacientes ainda que estejam sendo tratados com hidratação e indução de diurese. Os danos renais podem ocasionar hipomagnesemia,

hipocalemia, hipocalcemia, hipouricemia e hiponatremia que podem persistir por anos mesmo com a interrupção do tratamento com cisplatina (HARTMANN et al., 1999; FUKASAWA et al., 2014).

A cisplatina é um medicamento de ação intracelular, portanto, o primeiro passo, seja para executar sua ação terapêutica ou na indução de toxicidade, é a entrada nas células. Por ser uma molécula de baixo peso molecular (300,06 g/mol) e não carregada, a cisplatina passa livremente pela filtração glomerular, podendo ser recuperada na urina sem que tenha sofrido qualquer tipo de biotransformação. A excreção da cisplatina é predominantemente renal sendo acumulada nos rins durante o processo. Por isso, mesmo quando os níveis sanguíneos são seguros, os níveis renais podem ser tóxicos (SAFIRSTEIN; DAYE; GUTTENPLAN, 1983; PERES; CUNHA JÚNIOR, 2013).

Um dos transportadores de membrana responsável pela internalização da cisplatina é o transportador de cátion orgânico OCT2 (*SLC22A2*), membro da família de transportadores SLC22, expresso na membrana basolateral de células do túbulo contorcido proximal. OCTs possuem expressão órgão-específica, sendo que OCTs predominantemente expressos em outros órgãos não interagem com a cisplatina, o que torna o OCT2 um dos responsáveis pela especificidade da toxicidade renal. Foi observado que em camundongos OCT1 também é expresso nos rins, porém em nível menor que OCT2. Em humanos, OCT1 é predominantemente expresso nas células hepáticas (YONEZAWA et al., 2005; FILIPSKI et al., 2009; CIARIMBOLI et al., 2010; CIARIMBOLI, 2014).

O transportador de cobre CTR1 (*SLC31A1*) também participa da entrada da cisplatina nas células e localiza-se principalmente na membrana basolateral das células dos túbulos proximal e distal. No entanto, CTR1 é expresso em diversos órgãos além dos rins (ISHIDA et al., 2002; PABLA et al., 2009). Por isso, ainda que sua expressão renal seja importante na nefrotoxicidade induzida por cisplatina, esse transportador não é responsável pela especificidade da toxicidade renal. Na face apical das células do túbulo contorcido proximal está o transportador de extrusão de múltiplas drogas e toxinas MATE1 (*SLC47A1*), do qual a cisplatina também é substrato. Ao contrário dos transportadores supracitados, MATE1 está relacionado com a saída da cisplatina do interior da célula para o lúmen tubular, onde será excretada com a urina (YONEZAWA et al., 2006; NAKAMURA et al., 2010). Outros transportadores estão envolvidos no transporte de cisplatina em tumores, estando aqui destacados aqueles de importância para a nefrotoxicidade.

Uma vez dentro das células, o baixo nível de cloreto intracelular favorece que os átomos de cloro da cisplatina sejam substituídos por moléculas de água, sem que as

modificações químicas afetem os grupos amina ligados à platina. Esta modificação ativa a cisplatina, tornando-a carregada positivamente e apta a ligar-se ao DNA nuclear e mitocondrial e também ao RNA e proteínas. As espécies ativadas de cisplatina formam ligações covalentes cruzadas intrafita com o átomo de nitrogênio 7 (N7) de purinas (guanina e adenina), sendo o principal tipo de aduto cisplatina-DNA aquele formado entre guaninas adjacentes (1,2-d(GpG) intrafita). Em proporção muito menor formam-se adutos intrafita entre adenina e guanina adjacentes (1,2-d(ApG)). Adutos 1,3-intra e interfita também podem ocorrer, além de adutos entre cisplatina-DNA-proteína (Figura 4) (FICHTINGER-SCHEPMAN et al., 1985; DAVIES; BERNERS-PRICE; HAMBLEY, 2000; DASARI; TCHOUNWOU, 2014).

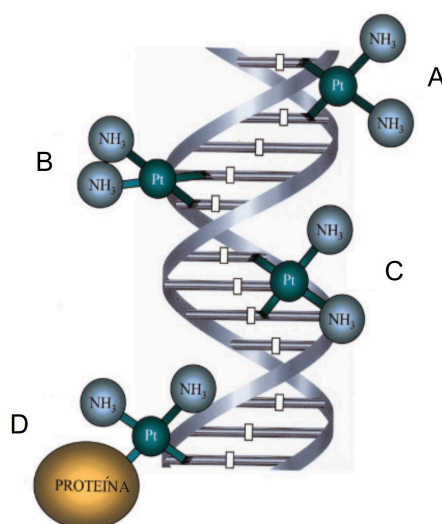


Figura 4. Tipos de adutos formados entre cisplatina e DNA. (A) Ligação cruzada interfita entre bases adjacentes. (B) Ligação cruzada intrafita entre bases adjacentes. (C) Ligação cruzada intrafita entre bases intercaladas. (D) Ligação cruzada entre DNA e proteína.

Adaptado de GONZALES et al. Is cisplatin-induced cell death always produced by apoptosis? **Molecular Pharmacology**, v. 59, n. 4, p. 657-663, 2001.

A formação do aduto cisplatina-DNA inibe os processos de replicação e transcrição, leva à parada do ciclo celular, induz disfunção mitocondrial, estresse do retículo endoplasmático, estresse oxidativo, inflamação e apoptose. As mitocôndrias possuem acúmulo de carga negativa o que as torna importantes alvos da cisplatina, que se torna positivamente carregada ao entrar na célula, como citado anteriormente. A concentração

celular de mitocôndrias é um fator determinante nos efeitos tóxicos da cisplatina (QIAN et al., 2005).

Dentre as células que compõem as diferentes regiões do néfron, aquelas do túbulo proximal são as que possuem maior concentração mitocondrial, o que ajuda a explicar o fato desta ser a região mais afetada pela cisplatina. A disfunção mitocondrial resulta na diminuição da síntese de ATP, levando a célula a uma condição de estresse que culmina na apoptose (morte celular programada). Nessas condições, as proteínas pró-apoptóticas BAX e BAK sofrem modificações estruturais que alteram a integridade da membrana mitocondrial. Dessa maneira, o citocromo c é liberado do interior da mitocôndria para o citoplasma, onde formará um complexo responsável pela ativação da caspase-9. Uma vez ativada, a caspase-9 ativará caspases-3, responsável pela execução da apoptose (KAUSHAL et al., 2001; SERVAIS et al., 2008).

A disfunção mitocondrial promove um aumento na produção de espécies reativas de oxigênio, levando a célula à condição de estresse oxidativo (PABLA; DONG, 2008; MANOHAR; LEUNG, 2018). A mitocôndria não é a única organela a sofrer alterações devido à exposição à cisplatina. O estresse do retículo endoplasmático induz apoptose através da ativação da caspase-12, uma proteína que se localiza no retículo endoplasmático e é liberada para o citoplasma devido a alterações na homeostase de cálcio e pelo acúmulo de proteínas no retículo endoplasmático. Ao ser liberada para o citoplasma, a caspase-12, assim como a caspase-9, ativa caspases responsáveis pela execução da apoptose (MANDIC et al., 2003).

Além da indução das vias intrínsecas da apoptose mencionadas acima, mitocondrial e do retículo endoplasmático, a cisplatina também induz a apoptose pela via extrínseca. A via extrínseca da apoptose é também conhecida como via dos receptores de morte por ser ativada pela interação do fator de necrose tumoral alfa ($TNF\alpha$) ao receptor TNFR. A interação entre ligante e receptor culmina na ativação da caspase-8 que, por sua vez, ativará caspases executoras da apoptose (SERVAIS et al., 2008). Adutos cisplatina-DNA causam quebra de fita simples e dupla de DNA. O dano causado pela cisplatina ao DNA ativa a via do supressor de tumor p53, que regula a progressão do ciclo celular e a apoptose, como forma de garantir a integridade do genoma. P53 é ativado pela proteína ATR e então forma um tetrâmero que se liga ao DNA. Este complexo atua induzindo ou reprimindo a transcrição de diversos genes que, em conjunto, atuarão no processo a ser regulado. P53 induz a apoptose através da transcrição de genes como *PUMA- α* e *PIDD*, o que culmina na liberação de citocromo c e AIF (do inglês, *apoptosis-inducing factor*), respectivamente, pela mitocôndria. AIF induz a

apoptose por uma via independente de caspases (SETH et al., 2005; JIANG et al., 2006; PABLA et al., 2008; RILEY et al., 2008).

A inflamação também é uma resposta decorrente do dano renal ocasionado por cisplatina e que também contribui para o dano. O estresse oxidativo ocasionado pelo aumento da produção de espécies reativas de oxigênio leva à ativação de NF- κ B. Esse último processo regula a expressão de muitos genes envolvidos na resposta inflamatória, dentre eles o TNF α , considerado importante na nefrotoxicidade induzida por cisplatina. O bloqueio de TNF α previne que outras citocinas sejam expressas como, por exemplo, TGF- β , IL-1 β , RANTES, MIP-2, MCP-1. Ainda que TNF α seja importante na inflamação durante a nefrotoxicidade induzida por cisplatina, outros componentes também participam do processo inflamatório, como é o caso de ICAM-1 (BARNES; KARIN, 1997; RAMESH; REEVES, 2002).

Pouco se sabe sobre o papel dos RNA não codificadores (ncRNA) na nefrotoxicidade induzida por cisplatina. No entanto, sabe-se que ncRNAs podem ser utilizados como biomarcadores e como alvos terapêuticos. Os estudos realizados até o momento focam principalmente nos microRNAs (miRNA) e evidências foram encontradas de que estão envolvidos na disfunção mitocondrial (GUO et al., 2018), estresse oxidativo (LIAO et al., 2017) e regulação da apoptose (BHATT et al., 2010). Até o presente momento, não foram identificados estudos que relacionam os RNAs longos não codificadores (lncRNA) diretamente à nefrotoxicidade induzida por cisplatina.

1.4 Transcriptômica e tecnologias associadas

O termo transcriptoma foi utilizado pela primeira vez no final da década de 1990 e pode ser definido como o conjunto de genes expressos em uma célula, tecido, órgão ou organismo em uma determinada situação. Entretanto, naquele momento nenhum transcriptoma havia sido caracterizado devido a limitações tecnológicas (VELCULESCU et al., 1997). Ainda que não fosse possível analisar o transcriptoma completo, análises de transcritos individuais eram realizadas graças ao método de sequenciamento de Sanger, desenvolvido em 1977. Com isso, foi possível identificar genes a partir do sequenciamento parcial de moléculas de DNA complementar (cDNA), sendo geradas etiquetas de sequências expressas (EST). Essa metodologia foi amplamente utilizada durante a década de 1990 (MARRA; HILLIER; WATERSTON, 1998).

A fim de possibilitar a quantificação dos transcritos e aumentar o rendimento das reações, a metodologia EST foi adaptada para novos métodos denominados análise serial ou de *cap* de expressão gênica (SAGE/CAGE) (VELCULESCU et al., 1995; SHIRAKI et al., 2003). Em meados da década de 1990, foi desenvolvida uma nova metodologia de identificação e quantificação de transcritos, denominada microarranjo, que se baseia na hibridização dos transcritos em sondas complementares fixadas em lâmina (SCHENA et al., 1995). Os microarranjos, diferentemente das metodologias utilizadas até então, permitem que milhares de transcritos sejam analisados simultaneamente (HELLER, 2002). Nos anos 2000, surgiu uma nova metodologia que modificou significativamente a forma com que os transcriptomas eram analisados, o sequenciamento de RNA de alto rendimento (RNA-Seq) (WANG; GERSTEIN; SNYDER, 2009).

O RNA-Seq passou a ser a metodologia mais empregada no estudo do transcriptoma a partir do final da década de 2000 (Figura 5). Dentre as vantagens do RNA-Seq estão o alto rendimento, o baixo custo e, principalmente, a possibilidade de identificar e quantificar transcritos sem o conhecimento prévio de sua sequência (LOWE et al., 2017). O uso mais frequente do RNA-Seq é a análise de expressão diferencial de genes, contudo outras aplicações vêm ampliando a compreensão de aspectos biológicos como é o caso da identificação de isoformas derivadas de *splicing* alternativo (WANG et al., 2008b), regulação da expressão gênica por ncRNAs (DJEBALI et al., 2012; MORRIS; MATTICK, 2014) e por RNAs potenciadores (*enhancer RNA*) (LI; NOTANI; ROSENFELD, 2016).

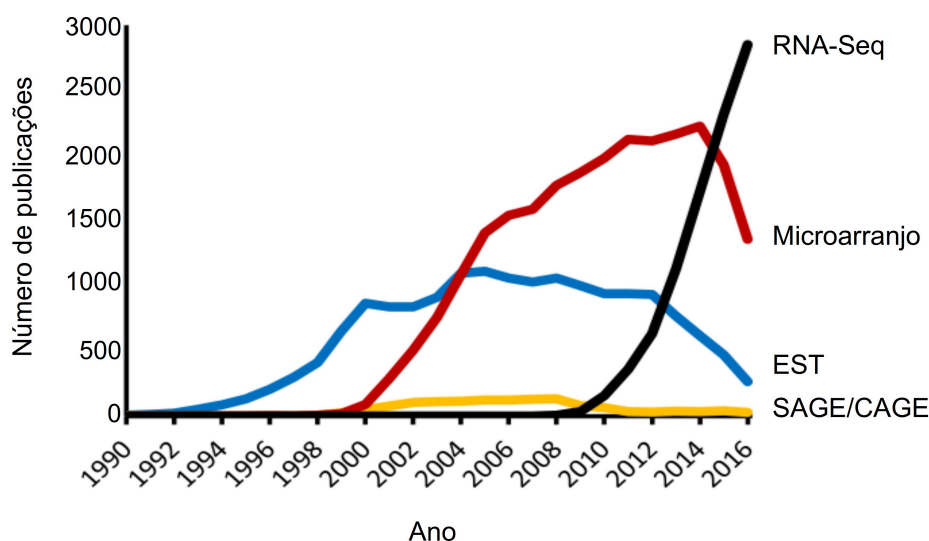


Figura 5. Tecnologias para estudo do transcriptoma no decorrer dos anos. Publicações científicas entre os anos de 1990 e 2016 referentes a RNA-Seq (linha preta), microarranjo (linha vermelha), EST (linha azul) e SAGE/CAGE (linha amarela). RNA-Seq, sequenciamento de RNA. EST, etiquetas de sequências expressas. SAGE/CAGE, análise serial/de *cap* de expressão gênica.

Adaptado de LOWE, R. et al. Transcriptomics technologies. **Plos Computational Biology**, v. 13, n. 5, p. e1005457, 2017.

Para a realização do sequenciamento, o RNA extraído precisa passar por processamentos, passo conhecido como construção da biblioteca de RNA. Resumidamente, é necessário enriquecer os tipos de RNAs com os quais se deseja trabalhar, o que pode ser feito por captura do RNA pela cauda poliA ou pela depleção dos RNAs ribossomais da amostra, no caso de transcritos de RNA polimerase II (Pol II). Na sequência, os RNAs são convertidos em cDNA, que são fragmentados e inseridas sequências adaptadoras nas extremidades desses fragmentos. Após a construção das bibliotecas, os fragmentos são aplicados na *flowcell* (lâmina utilizada pelo sequenciador) e os adaptadores hibridizarão com sequências contidas na *flowcell*. Os fragmentos são amplificados por PCR e, então, é realizado o sequenciamento em si. Os passos acima descritos são utilizados para o sequenciamento por síntese de *reads* curtas, tecnologia utilizada nos sequenciamentos da empresa Illumina (Figura 6) (LOWE et al., 2017; STARK; GRZELAK; HADFIELD, 2019).

O sequenciamento pode ser realizado a partir de uma extremidade do fragmento (*single-end*) ou a partir das duas extremidades (*paired-end*). Além disso, é possível que sejam inseridas sequências com função de códigos de barra (*barcodes*) que são utilizadas para a posterior identificação da fita de DNA que originou o fragmento (sequenciamento *strand-specific*) durante o preparo da biblioteca. O processamento dos dados sequenciados demanda

grande poder computacional e envolve os passos de alinhamento contra o genoma ou transcriptoma de referência (ou montagem do transcriptoma) que permite a identificação dos transcritos, quantificação e análise da expressão dos mesmos (LOWE et al., 2017; STARK; GRZELAK; HADFIELD, 2019).

O sequenciamento por síntese de *reads* curtas (Illumina) é o principal método aplicado para geração de dados de RNA-Seq disponíveis na base de dados SRA (*Sequence Read Archive*) (LEINONEN; SUGAWARA; SHUMWAY, 2011), embora outras metodologias estejam em constante desenvolvimento. Atualmente, metodologias capazes de sequenciar *reads* longas de cDNA (Pacific Biosciences) ou realizar o sequenciamento direto de RNA (Oxford Nanopore) já estão disponíveis (Figura 6).

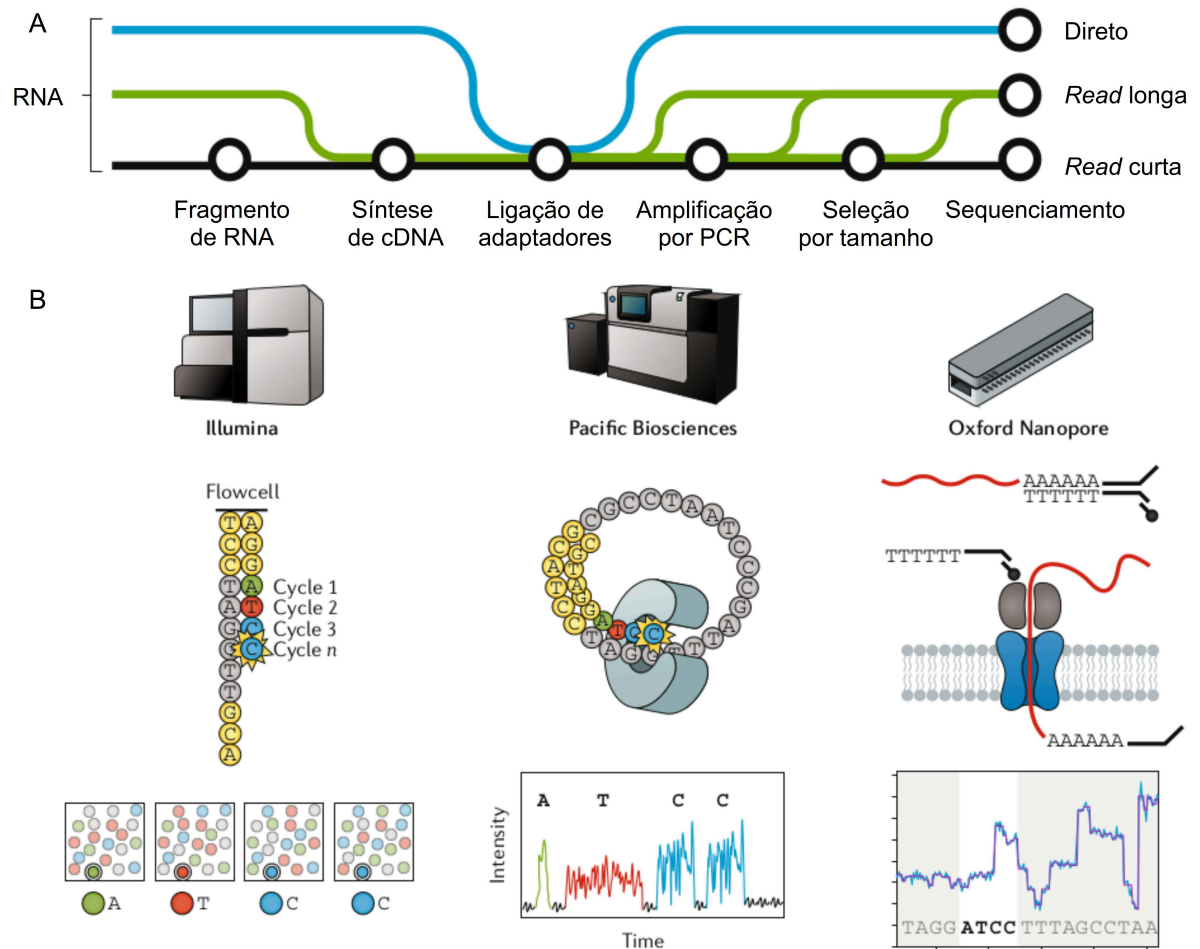


Figura 6. Tecnologias de sequenciamento de RNA. (A) Visão geral dos métodos de preparação de bibliotecas para diferentes métodos de sequenciamento de RNA classificados como sequenciamento de cDNA de *reads* curtas (linha preta), sequenciamento de cDNA de *reads* longas (linha verde) ou sequenciamento direto de *reads* longas (linha azul). (B) Visão geral de três principais tecnologias de sequenciamento de RNA. À esquerda, tecnologia Illumina, onde moléculas individuais de cDNA são agrupadas em uma *flowcell* para sequenciamento por síntese utilizando nucleotídeos marcados com fluorescência, a cada ciclo de incorporação é detectado qual nucleotídeo foi adicionado à cadeia. Essa metodologia é capaz de sequenciar cDNA de *reads* curtas. Ao meio, tecnologia Pacific Biosciences, onde moléculas individuais são carregadas em um chip de sequenciamento, onde se ligam a uma polimerase imobilizada no fundo de um nanoporo. Nucleotídeos marcados com fluorescência são incorporados à fita em crescimento e são detectados. Essa metodologia sequencia cDNA de *reads* longas. À direita, tecnologia Oxford Nanopore, onde moléculas individuais são carregadas em uma *flowcell*, onde proteínas motoras, que são conectadas durante a ligação do adaptador, acoplam-se ao nanoporo. As proteínas motoras controlam a translocação dos RNAs através dos nanoporos, causando uma mudança na corrente elétrica que é processada para identificação dos nucleotídeos. Essa metodologia faz sequenciamento direto de *reads* longas. cDNA, DNA complementar.

Adaptado de STARK, R., GRZELAK, M., HADFIELD, J. RNA sequencing: the teenage years. *Nature Reviews Genetics*, v. 20, p. 631-656, 2019.

1.5 Toxicogenômica

A toxicologia estuda os efeitos adversos causados por substâncias diversas em organismos vivos. Para isso, a toxicologia tradicional avalia efeitos citológicos, fisiológicos, metabólicos e morfológicos. Em geral, cumprir a carga de testes toxicológicos demanda muito tempo, custos e trabalho (SUTER; BABISS; WHEELDON, 2004). Tendo em vista que a expressão gênica é modificada em face à toxicidade, a toxicogenômica surgiu no final da década de 1990 com o advento do microarranjo, sendo vista como uma forma de revolucionar os estudos toxicológicos (NUWAYSIR et al., 1999).

A toxicogenômica é definida como a combinação da toxicologia com tecnologias -ômicas de alto rendimento, como genômica, transcriptômica, proteômica e metabolômica, para estudar os efeitos adversos de produtos químicos, ambientais e farmacêuticos na saúde humana e no meio ambiente através de alterações específicas nos padrões de expressão de genes, proteínas e metabólicos com respostas fenotípicas de células, tecidos e organismos (NATIONAL RESEARCH COUNCIL, 2007).

A toxicologia mecanística busca identificar os meios pelos quais ocorre a toxicidade em nível celular e tecidual, definindo as bases bioquímicas e fisiológicas envolvidas no processo (BARILE, 2010). Esta área da toxicologia tem enorme potencial para beneficiar-se da toxicogenômica, pois ao avaliar e caracterizar a expressão diferencial de genes após a exposição a xenobióticos, é possível utilizá-los para prever resultados toxicológicos e identificar mecanismos envolvidos ou relacionados ao evento tóxico (SUTER; BABISS; WHEELDON, 2004).

Estudos toxicogenômicos baseiam-se amplamente na tecnologia do microarranjo. Há um questionamento em aberto sobre a migração das abordagens baseadas em microarranjos para o RNA-Seq, devido ao histórico de duas décadas de acúmulo de informações geradas com microarranjos e a maturidade de análise dentro do campo da toxicologia (LIU et al., 2019b; MEZENCEV; SUBRAMANIAM, 2019).

Em contrapartida, sabe-se que o RNA-Seq apresenta vantagens na detecção de genes de baixa expressão, identificação de novos transcritos derivados de *splicing*, sítio de início de transcrição e poliadenilação alternativos (SEQC/MAQC-III CONSORTIUM, 2014), além das vantagens já citadas no tópico “1.4 Transcriptômica e tecnologias associadas”. A maior precisão na detecção de transcritos de baixa abundância tem mostrado a superioridade do RNA-Seq aos microarranjos na detecção de genes diferencialmente expressos em estudos toxicológicos (WANG et al., 2014a; RAO et al., 2019). Isso, aliado às menores limitações ao

longo do tempo, tendem a tornar o uso do RNA-Seq mais frequente, especialmente em projetos que precisam quantificar transcritos de baixa abundância, variantes de *splicing* alternativo e ncRNAs (MEZENCEV; SUBRAMANIAM, 2019).

1.6 Genes não codificadores de proteínas

Com o surgimento das tecnologias de sequenciamento de DNA e RNA de alto rendimento e profundidade, descobriu-se que o genoma humano, de camundongo e de diversos eucariotos não são compostos majoritariamente de genes codificadores de proteínas (MEISTER, 2011). O genoma humano é composto por mais de 60.500 genes, destes, menos de 20.000 referem-se a genes codificadores de proteínas e mais de 25.000 são classificados como genes não codificadores de proteínas. Já no genoma do camundongo foram identificados mais de 50.000 genes, sendo aproximadamente 20.000 genes codificadores de proteínas e um número similar de genes não codificadores de proteínas (GENCODE, 2019a, 2019b). Em animais, a complexidade biológica não se correlaciona com o número de genes codificadores de proteínas. Ao passo que existe uma correlação forte entre a proporção não codificadora do genoma e a complexidade do organismo (LIU; MATTICK; TAFT, 2013).

Os genes codificadores de proteínas são aqueles que contêm uma sequência codificadora (CDS, do inglês *coding sequence*). Já os genes não codificadores não possuem CDS e são denominados, genericamente, de genes de transcritos processados, sendo divididos em três grandes categorias de acordo com a classificação do Ensembl, sendo elas: lncRNA, pequenos RNAs não codificadores e transcritos processados não classificados. Os genes de lncRNA podem ser subdivididos de acordo com suas localizações genômicas, tendo como característica comum o fato de possuírem mais que 200 pares de bases (pb). Dentre os lncRNAs estão as categorias antissenso, que são genes localizados na fita oposta, sobrepostos, a genes codificadores e os RNAs longos não codificadores intercalados ou intergênicos (lincRNA, do inglês *long interspersed ncRNA*). Os genes também podem ser classificados como pseudogenes, que são sequências derivadas dos genes codificadores de proteínas que perderam o potencial codificador devido a inserções e deleções que modificam a fase de leitura ou por mutações que inserem códons de parada precoces (ENSEMBL, [201-]; FRANKISH et al., 2019). Os lncRNAs são transcritos pela Pol II, assim como os RNA mensageiros (mRNA), em sua maioria sofrem *splicing* alternativo e também podem ser poliadenilados. Apenas uma pequena fração, entre 11% e 29%, dos lncRNAs são ubiquamente

expressos em todos os tecidos, ao passo que mais de 60% dos genes codificadores de proteína o são. Ainda em comparação aos genes codificadores de proteína, os lncRNAs tendem a ter uma expressão muito mais baixa (DERRIEN et al., 2012; DJEBALI et al., 2012).

Os lncRNAs participam da regulação da expressão gênica em seus diversos níveis, desde o transcricional até o traducional (Figura 7). Ao nível transcricional, os lncRNAs participam de modificações epigenéticas de histonas, alterando a estrutura da cromatina. Com isso, os genes localizados nessas regiões podem ter sua expressão ativada ou suprimida (GUPTA et al., 2010; WANG et al., 2011). LncRNAs também podem atuar na regulação da transcrição, ligando-se a fatores de transcrição. Através desse mecanismo, a expressão pode ser ativada quando o lncRNA recruta os fatores de transcrição para seus devidos promotores ou inibida se o lncRNA atua sequestrando os fatores de transcrição (MARTENS; LAPRADE; WINSTON, 2004; WILLINGHAM et al., 2005). Já o *splicing* alternativo pode ser modulado por lncRNAs através da modificação do sítio de *splicing* a ser utilizado (HASTINGS et al., 1997) e também pela interação com fatores de *splicing* (HE; LUO; MO, 2019).

Na regulação ao nível traducional, os lncRNAs participam do controle de expressão atuando como esponjas de miRNAs e, assim, inibindo sua função (CESANA et al., 2011; HANSEN et al., 2013), estabilizando mRNAs e promovendo a tradução (CARRIERI et al., 2012) ou a interrompendo (EBRALIDZE et al., 2008). Os lncRNA também participam do decaimento de mRNAs localizando proteínas que marcam o mRNA para degradação (GONG; MAQUAT, 2011).

Em processos fisiológicos, os lncRNAs estão envolvidos, por exemplo, na embriogênese na manutenção da pluripotência, proliferação e migração celular e apoptose (LAURENT et al., 2016). Os lncRNAs também já foram associados a diversas doenças, tais como doenças cardíacas (LI; ZHOU; HUANG, 2017; ZHANG et al., 2018b), doença de Alzheimer (MAGISTRI et al., 2015), doença celíaca (CASTELLANOS-RUBIO et al., 2016) e câncer (BLUME et al., 2015; CHEN et al., 2017; KONG; QIU, 2018). Adicionalmente, estão envolvidos na resposta tóxica devido à exposição a substâncias como benzeno (BAI et al., 2014), cádmio (TANI et al., 2014), níquel (ZHANG et al., 2012) e ftalatos (LAROCCA et al., 2014).

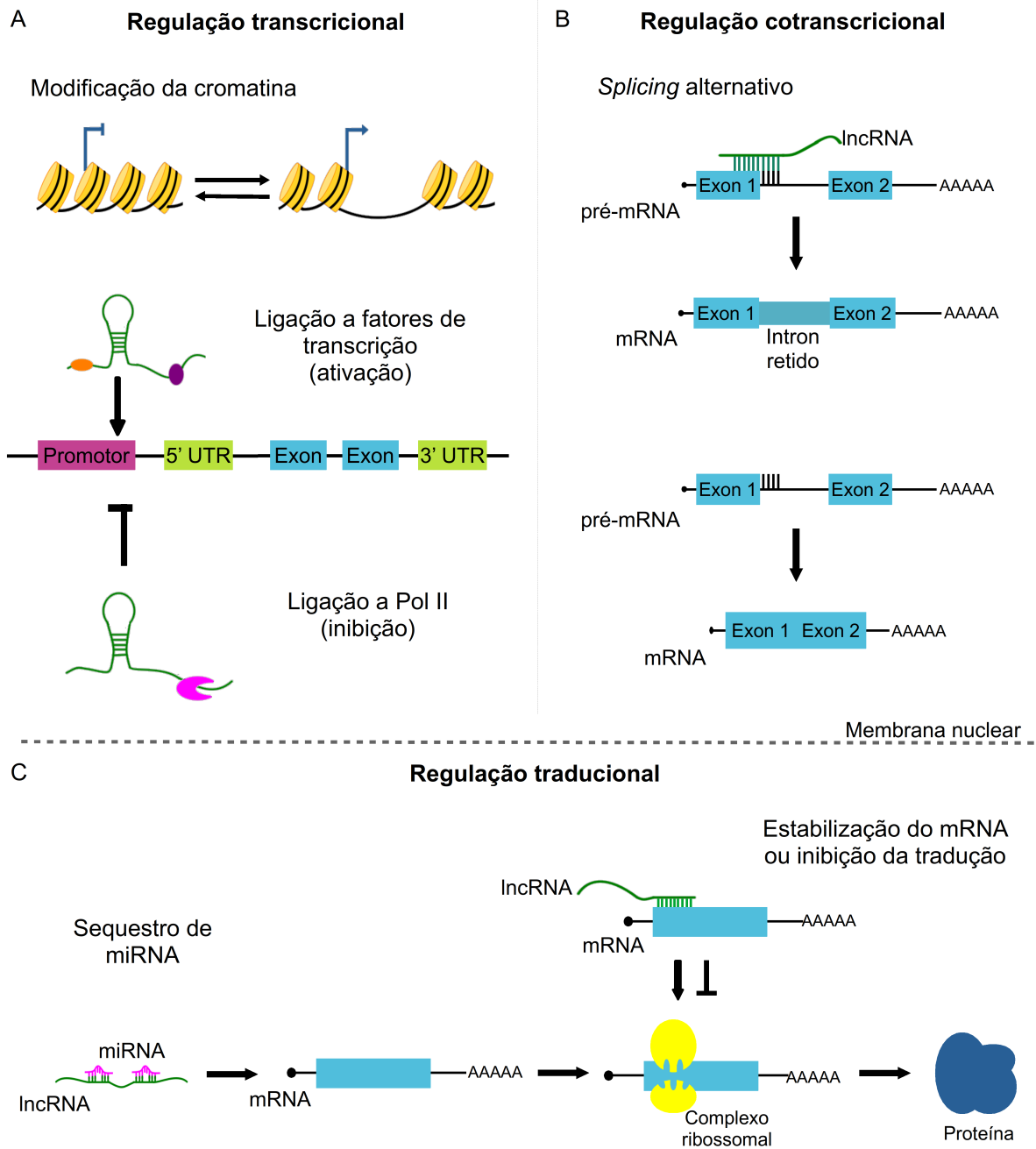


Figura 7. Exemplos de funções dos lncRNAs. (A) Regulação ao nível transcricional. Acima, modulação da estrutura da cromatina através da interação de lncRNA com modificadores de histonas. Abaixo, lncRNAs influenciam a transcrição através do recrutamento e localização de fatores de transcrição ou a inibem através do sequestro de Pol II. (B) Regulação do *splicing* alternativo. Acima, a ligação de um lncRNA ao mRNA promove a retenção de íntron. Abaixo, a ausência do lncRNA não altera o processamento e o íntron é retirado durante o *splicing*. (C) Regulação ao nível da tradução. À esquerda, lncRNAs sequestram miRNAs. À direita, lncRNAs interagem com mRNA aumentando a estabilidade do transcrito ou inibindo a tradução. lncRNA, RNA longo não codificador. mRNA, RNA mensageiro. miRNA, microRNA. Pol II, RNA polimerase II.

Adaptado de DEMPSEY, J.L., CUI, J.Y. Long non-coding RNAs: A novel paradigm for toxicology. *Toxicological Sciences*, v. 115, p. 3-21, 2017.

Ainda existem poucos estudos sobre o papel dos lncRNA na LRA (BRANDENBURGER et al., 2018). Tendo em vista que a mesma pode ser ocasionada por

diferentes fatores, diversos modelos são utilizados, dentre eles estão a hipóxia e a sepse. Yu e colaboradores (2016) (YU et al., 2016) observaram que o lncRNA PRINS (*psoriasis susceptibility-related RNA gene induced by stress*) tem papel na infiltração e ativação de células inflamatórias em resposta à hipóxia. A resposta inflamatória também é modulada pelo lncRNA PVT1 (*plasmacytoma variant translocation 1*) (HUANG et al., 2017) e por NEAT1 (*nuclear enriched abundant transcript 1*) (CHEN et al., 2018a) após LRA séptica induzida por lipopolissacarídeo. Ainda há muito a ser elucidado sobre o papel dos lncRNAs na nefrotoxicidade induzida por cisplatina e na LRA. O fato de lncRNAs participarem da regulação de processos sabidamente envolvidos na nefrotoxicidade induzida por cisplatina e LRA, tais como apoptose, controle do ciclo celular (HUARTE et al., 2010; WU et al., 2014), inflamação e estresse oxidativo (LI; ZHOU; HUANG, 2017; XIAO et al., 2019) constitui evidência de que os lncRNAs têm importante papel em ambas as condições.

1.7 *Splicing* alternativo como mecanismo de variabilidade de transcritos

Independentemente do status de codificador ou não codificador de proteínas, os transcritos produzidos pela Pol II podem ser processados diferentemente ampliando o repertório de RNAs derivados de um mesmo gene. Como citado anteriormente, os genomas humano e do camundongo possuem mais de 60.500 e 50.000 genes, respectivamente. Entretanto, o número de transcritos gerados por esses genes é de mais de 200.000 em humanos e, aproximadamente, 140.000 em camundongos (GENCODE, 2019a, 2019b). Essa grande diferença entre o número de genes e o número de transcritos se deve, em parte, ao mecanismo de *splicing* alternativo. O *splicing* é o processamento do RNA em que os íntrons são removidos e os éxons são unidos gerando transcritos maduros. No *splicing* alternativo, diferentes éxons e íntrons de um mesmo RNA são mantidos ou excluídos para gerar transcritos maduros ligeiramente distintos, aumentando a diversidade do transcriptoma e também do proteoma (KEREN; LEV-MAOR; AST, 2010).

O processo de *splicing* é realizado pelo spliceossomo, que é um complexo ribonucleoproteico composto por cinco pequenas ribonucleoproteínas associadas a um grande número de proteínas acessórias que reconhecem os sítios de *splicing* nas extremidades de éxons e íntrons, realizam a remoção dos íntrons e ligam os éxons durante o processamento do RNA que está sendo transcrito (JURICA; MOORE, 2003). É chamado de *splicing* canônico aquele que mantém todos os éxons e de *splicing* alternativo aquele em que há algum tipo de

modificação na estrutura do transcrito. Dentre os tipos conhecidos de *splicing* alternativo, estão a exclusão de éxon, retenção de íntron, sítio de *splicing* 5' e 3' alternativos e éxons mutuamente exclusivos (Figura 8).

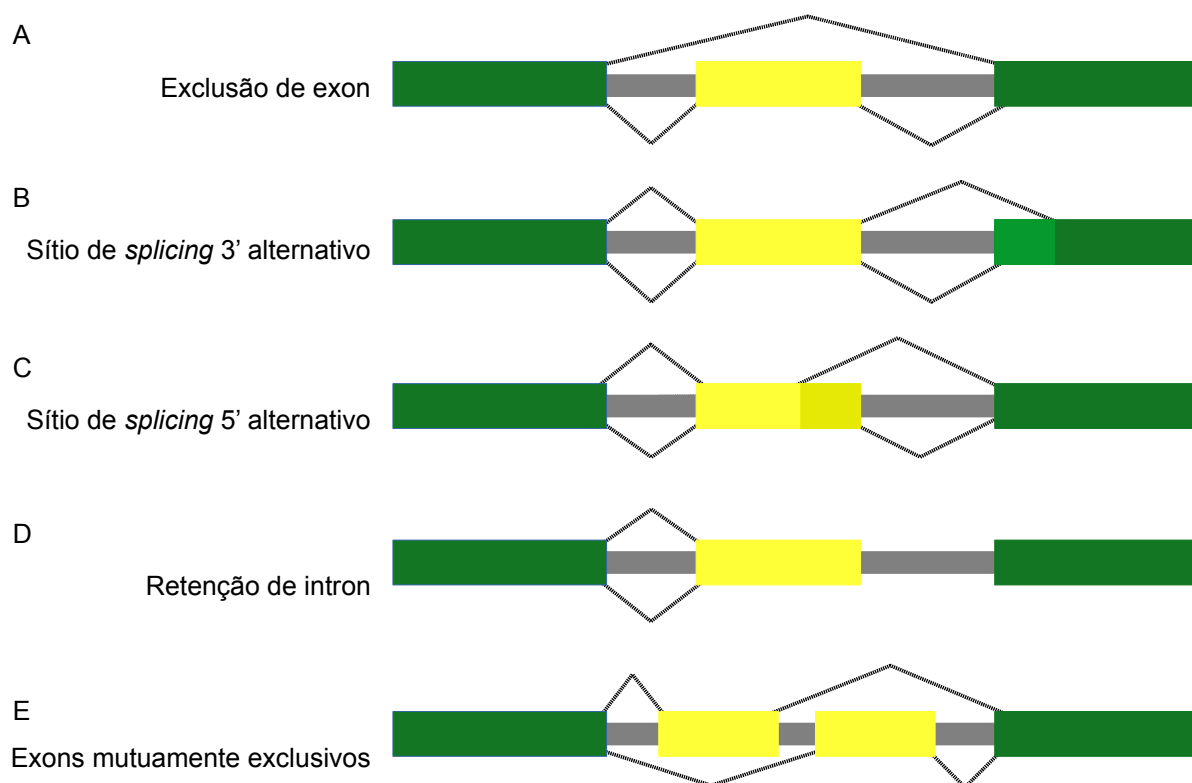


Figura 8. Diferentes tipos de *splicing* alternativo. (A) Exclusão de éxon, um ou mais éxons podem ser excluídos do transcrito maduro. (B) Sítio de *splicing* 3' alternativo, uso de um sítio alternativo de *splicing* à montante do sítio constitutivo de forma a modificar a estrutura do éxon. (C) Sítio de *splicing* 5' alternativo, uso de um sítio alternativo de *splicing* à jusante do sítio constitutivo de forma a modificar a estrutura do éxon. (D) Retenção de íntron, íntron não é removido durante o *splicing* e permanece no transcrito maduro. (E) Éxons mutuamente exclusivos, em um conjunto de dois éxons, um obrigatoriamente será excluído na presença do outro. Linhas pontilhadas representam as possibilidades de *splicing*, caixas amarelas e verdes representam os éxons e os íntrons estão representados em cinza.

A exclusão de éxon ocorre quando um ou mais éxons são removidos do transcrito maduro juntamente a seus íntrons flanqueadores (Figura 8A). A retenção de íntron ocorre quando um íntron é mantido no transcrito maduro (Figura 8D). O uso de sítios de *splicing* 3' ou 5' alternativos ocorre quando além do sítio de *splicing* constitutivo há outro sítio concorrente à jusante ou à montante do sítio canônico (Figuras 8B e 8C). O uso do sítio alternativo de *splicing* resulta em modificação da estrutura do éxon, fazendo com que o mesmo seja maior ou menor do que o éxon gerado pelo sítio constitutivo. Já no tipo de *splicing* alternativo denominado éxons mutuamente exclusivos, em um conjunto de dois ou

mais éxons, o processamento não ocorre de maneira independente. Dessa maneira, quando um éxon é mantido no transcrito maduro o outro é obrigatoriamente removido (Figura 8E) (MATLIN; CLARK; SMITH, 2005).

Tanto genes codificadores quanto genes não codificadores de proteínas podem sofrer *splicing* alternativo. Um ponto fundamental dos genes codificadores de proteínas é que o transcrito maduro pode permanecer sendo codificador de proteína ou pode perder essa capacidade dependendo da maneira que o *splicing* alternativo é regulado.

1.8 Transcritos originados de genes codificadores de proteínas

Assim como os genes, os transcritos também podem ser classificados em diferentes tipos. Neste tópico serão abordados os tipos de transcritos que podem ser gerados a partir de genes codificadores de proteínas decorrentes do *splicing* alternativo.

Os genes codificadores de proteínas são aqueles que possuem CDS. Já os transcritos podem ser classificados como codificadores de proteínas, NMD (do inglês, *nonsense mediated decay*) ou íntron retido. Os transcritos NMD são aqueles que possuem um códon de parada precoce e que levam esse transcrito a ser degradado pela via NMD, enquanto os transcritos classificados como íntron retido são aqueles que mantêm um íntron no transcrito maduro (FRANKISH et al., 2019). Adicionalmente, o Ensembl classifica como NMD transcritos que possuem códon de parada, no mínimo, 50 nt antes de um sítio de *splicing* e caracteriza como não codificadores de proteína aqueles transcritos classificados como íntron retido (ENSEMBL, [201-]).

A presença do códon de parada precoce nos transcritos NMD é reconhecida durante o processo de tradução e faz parte do controle de qualidade dos mRNAs. Esse mRNA, que codifica uma proteína truncada, é degradado para que não ocorra um acúmulo desses transcritos no citoplasma bem como de proteínas potencialmente danosas que seriam por ele codificadas (LYKKE-ANDERSEN; JENSEN, 2015). Transcritos de isoformas NMD participam da regulação da expressão gênica tanto em condições fisiológicas quanto em condições patológicas. Alguns fatores de *splicing* autorregulam sua expressão produzindo isoformas NMD (JUMAA; NIELSEN, 1997; SUREAU et al., 2001), além disso, a via do NMD é essencial durante a embriogênese de mamíferos (MEDGHALCHI et al., 2001; MCILWAIN et al., 2010). A desregulação da via do NMD está envolvida em doenças como esclerose lateral amiotrófica (KAMELGARN et al., 2018), no desenvolvimento de

esquizofrenia infantil, desordens do espectro autista e déficit de atenção e hiperatividade (ADDINGTON et al., 2011) e no câncer (LINDEBOOM; SUPEK; LEHNER, 2016).

A retenção de íntron atua como um mecanismo de regulação da expressão gênica reduzindo níveis de transcritos desnecessários à fisiologia celular. A retenção de íntron pode levar à inserção de códon de parada precoce, gerando isoformas NMD que são exportadas para o citoplasma. Entretanto, a retenção de íntron pode levar à formação de isoformas não NMD que possuem níveis nucleares muito maiores que citoplasmáticos (BRAUNSCHWEIG et al., 2014). A retenção de íntron está envolvida em diversas doenças como câncer (ESWARAN et al., 2013; DVINGE; BRADLEY, 2015), xeroderma pigmentoso (CARTAULT et al., 2011) e doença inflamatória intestinal (HÄSLER et al., 2011).

1.9 Transcrição *readthrough* – transcrição de término estendido

O término da transcrição em eucariotos pela Pol II é um processo complexo que envolve inúmeras etapas (Figura 9). Resumidamente, após a sequência consenso de terminação da transcrição AAUAAA ser transcrita, proteínas de reconhecimento se ligam à mesma no RNA nascente. Em seguida, o RNA é clivado e inicia-se a síntese da cauda poliA pela enzima poli-A-polimerase (PAP) que finaliza o processamento da extremidade 3' do novo RNA. As proteínas de reconhecimento da sequência AAUAAA fator de estimulação da clivagem (CstF, do inglês *cleavage stimulation factor*) e fator de especificidade de clivagem e poliadenilação (CPSF, do inglês *cleavage and polyadenylation specificity factor*) movimentam-se com a cauda da Pol II e, com isso, ligam-se à sequência consenso logo após sua transcrição. Ao contrário da sequência consenso, a cauda poliA não é codificada no DNA. Assim, tem-se que a PAP é um tipo de RNA polimerase que não necessita de molde (ALBERTS et al., 2017).

Dois modelos, não mutuamente exclusivos, são propostos para liberação de Pol II do DNA ao término da transcrição. No modelo alostérico, Pol II se desliga do DNA molde devido a mudanças conformacionais desencadeadas pela clivagem do RNA. Já o modelo torpedo propõe que uma exonuclease, Xrn2 em mamíferos, degrada o RNA que continua sendo transcrito após a clivagem, a partir de sua extremidade 5' desprovida de *cap*. A degradação mais rápida que a síntese faz com que a exonuclease alcance Pol II e interaja com proteínas ligadas à sua cauda, desencadeando a liberação de Pol II do DNA molde (KREBS; KILPATRICK; GOLDSTEIN, 2014).

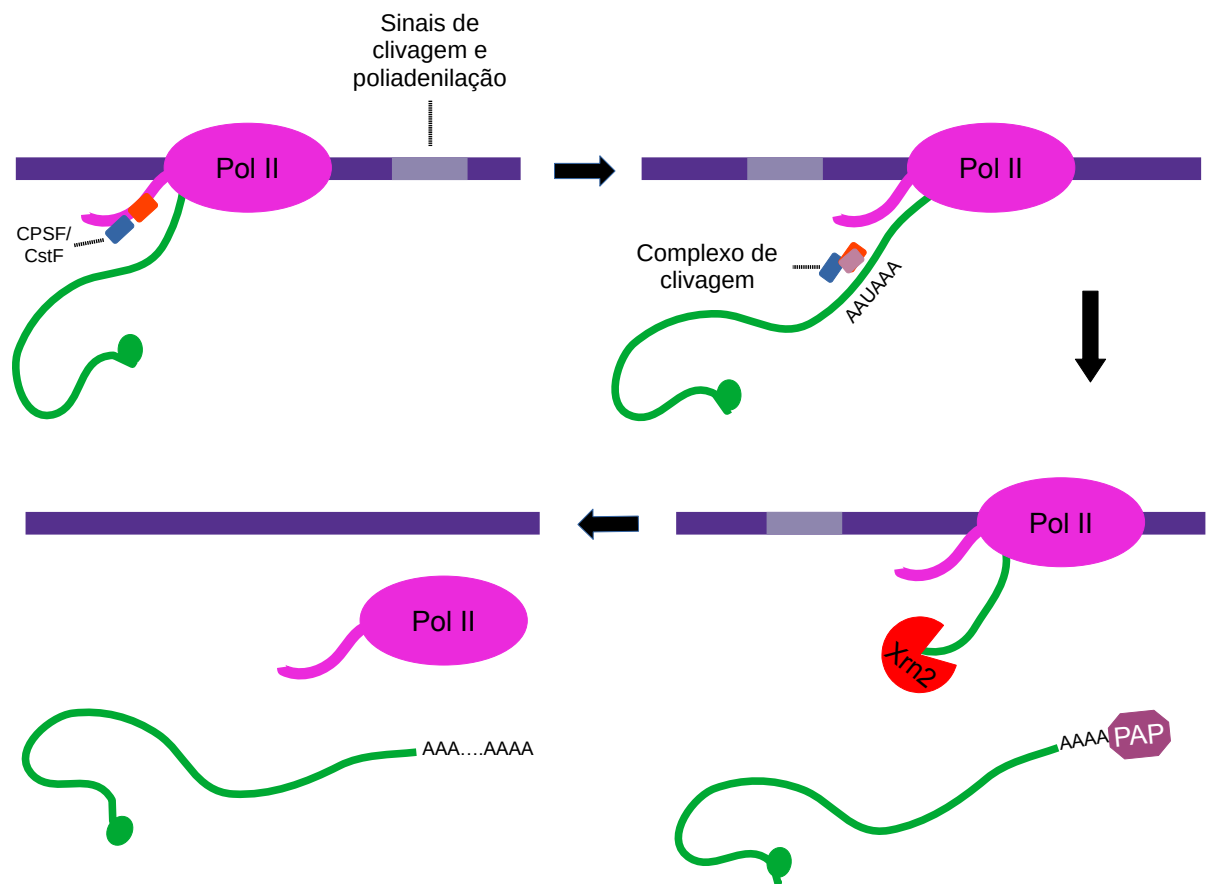


Figura 9. Principais etapas da terminação da transcrição e liberação de RNA polimerase II. Após a transcrição da sequência consenso de clivagem e poliadenilação (AAUAAA), proteínas de reconhecimento se ligam à sequência transcrita. O RNA nascente é clivado e a enzima poli-A-polimerase (PAP) sintetiza a cauda poliA, finalizando o processamento da extremidade 3' do novo RNA. A exonuclease Xrn2 degrada o RNA que continua sendo transcrito pela Pol II numa velocidade maior que sua síntese. A interação de Xrn2 com proteínas da cauda da Pol II induz a liberação da Pol II do DNA molde. Pol II, RNA polimerase II.

A falha em uma ou mais etapas do término da transcrição resulta na continuação da transcrição para além do sítio de terminação (Figura 10). Esse fenômeno, conhecido por transcrição *readthrough*, gera transcritos que têm se tornado conhecido por DoG (do inglês, *Downstream Of a Gene containing transcript*) (VILBORG; STEITZ, 2017). Em 2015, Vilborg e colaboradores (VILBORG et al., 2015) observaram que, sob estresse osmótico, DoGs podem ser responsáveis por até 20% da transcrição intergênica em linhagem celular de neuroblastoma humano SK-N-BE(2)C. Neste mesmo estudo, foi demonstrado que a transcrição de DoGs não é independente do gene à montante e que ambos fazem parte do mesmo transcrito. Além disso, demonstraram que a indução de DoGs parece estar relacionada à redução da terminação da transcrição do gene à montante e não ao aumento da transcrição desse gene.

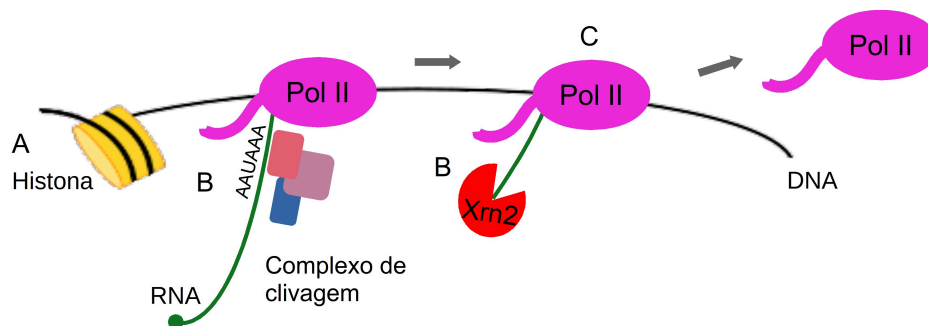


Figura 10. Fatores que podem afetar a eficiência de terminação da transcrição. (A) Modificações de cromatina que afetem a taxa de alongamento da Pol II. (B) Fatores envolvidos na clivagem e poliadenilação do transcrito. (C) Pol II pode afetar a eficiência de terminação da transcrição. Pol II, RNA polimerase II.

Adaptado de VILBORG, A., STEITZ, J.A. Readthrough transcription: how are DoGs made and what do they do? *RNA Biology*, v. 14, n. 5, p. 632-636, 2017.

Pouco se sabe a respeito dos mecanismos pelos quais os DoGs são gerados. O relaxamento da cromatina devido à hiperacetilação de histonas pode ser promovido através do aumento intracelular de cálcio, o que leva ao aumento da taxa de alongamento de Pol II (SHARMA et al., 2014). Sabendo que o processo de terminação da transcrição envolve uma pausa da Pol II próximo ao fim do gene, o aumento na taxa de alongamento da transcrição poderia afetar a terminação da mesma. Além disso, a sinalização de cálcio participa da ativação de diversas proteínas quinases e muitos fatores envolvidos na clivagem e poliadenilação são alvos de fosforilação (RYAN; BAUER, 2008). Assim, é possível que a sinalização de cálcio regule fatores envolvidos no término da transcrição ou a RNA Pol II de maneira a possibilitar a transcrição *readthrough*, gerando DoGs (VILBORG; STEITZ, 2017). De fato, sob condição de estresse osmótico, foi observado que a liberação de cálcio do retículo endoplasmático é necessária à indução de DoGs (VILBORG et al., 2015). Divergências na sequência consenso de poliadenilação também parecem estar envolvidas na transcrição de DoGs. Genes que possuem a sequência canônica AAUAAA parecem ser menos susceptíveis à transcrição *readthrough* em relação aos genes que possuem variações dessa sequência (VILBORG et al., 2015, 2017).

O papel dos DoGs ainda é pouco compreendido, entretanto, a indução de expressão de DoGs já foi observada na resposta à infecção pelo Vírus Herpes Simples (HSV-1) (RUTKOWSKI et al., 2015), na resposta celular ao estresse osmótico, onde os DoGs parecem estar relacionados à manutenção da integridade nuclear, em na resposta ao choque térmico (VILBORG et al., 2015) e em câncer renal, onde o número de genes com transcrição

readthrough correlaciona-se significativamente com um prognóstico ruim (GROSSO et al., 2015).

Apesar da nefrotoxicidade induzida por cisplatina ser conhecida desde os primeiros anos de teste deste medicamento e de constantes pesquisas com foco em desvendar mecanismos e propor alternativas para redução dos danos renais serem realizadas, os mecanismos de toxicidade ainda não estão completamente elucidados.

O crescente conhecimento sobre os processos de transcrição e de *splicing*, bem como a regulação dos mesmos, tem mostrado sua importância para a manutenção da homeostasia. A partir do processo de *splicing* alternativo, genes codificadores de proteínas podem dar origem tanto a transcritos codificadores quanto não codificadores. Adicionalmente, os genes não codificadores de proteína estão envolvidos em diversos mecanismos celulares, como demonstrado anteriormente. No entanto, pouco se sabe sobre a atuação de isoformas de transcritos e de genes não codificadores na nefrotoxicidade induzida por cisplatina. Assim, a análise de expressão gênica ao nível de transcritos pode revelar novos aspectos de mecanismos biológicos conhecidos e, com isso, abrir novas possibilidades de intervenção terapêutica.

Assim como a indução ou repressão de gene, o término da transcrição também deve ser corretamente coordenado. Problemas que afetam o término da transcrição e geram os transcritos denominados DoGs têm sido relacionados à situações de estresse celular. Sabendo que a cisplatina induz estresse devido ao dano de DNA, torna-se relevante verificar a existência de DoGs e seu padrão de expressão. Em conjunto, as abordagens propostas no presente trabalho têm como intuito ampliar os horizontes de conhecimento sobre os mecanismos de nefrotoxicidade.

2 OBJETIVOS

2.1 Objetivo geral

Elucidar mecanismos moleculares da nefrotoxicidade induzida por cisplatina através da análise da modulação da expressão gênica em nível de transcritos de genes codificadores e não codificadores de proteínas, incluindo suas várias isoformas, e da existência de eventos de transcrição *readthrough*.

2.2 Objetivos específicos

1. Analisar o efeito do tratamento na expressão diferencial de transcritos e os processos biológicos afetados pela cisplatina em rins de camundongos;
2. Investigar o papel de genes não codificadores de proteína na toxicidade renal induzida por cisplatina;
3. Avaliar a modificação, induzida pela cisplatina, dos padrões de *splicing* através da análise da expressão de transcritos não codificadores originados de genes codificadores de proteínas;
4. Avaliar a existência de eventos de transcrição *readthrough* e seu possível papel na nefrotoxicidade causada por cisplatina.

3 METODOLOGIA

3.1 Fluxo de trabalho resumido

O fluxo de trabalho resumido utilizado neste trabalho é apresentado na Figura 11.

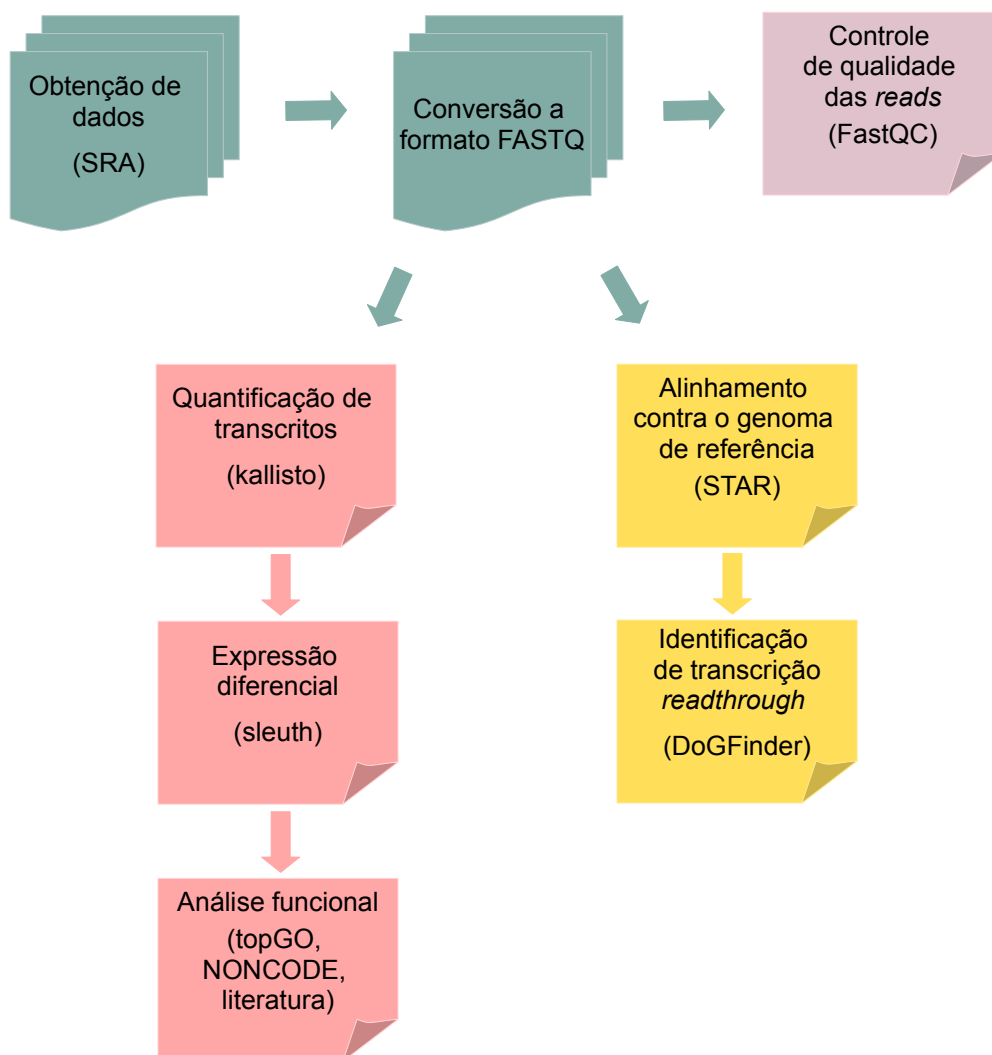


Figura 11. Fluxo de trabalho resumido.

3.2 Aquisição de dados de sequenciamento de RNA e controle de qualidade

Bibliotecas públicas de RNA-Seq disponíveis no banco de dados SRA mantido pelo *National Center for Biotechnology Information* (NCBI) foram utilizadas como fonte de informação para as análises realizadas neste trabalho. O SRA é uma base que armazena dados

de sequenciamento de alto rendimento oriundos de diversas plataformas (NCBI, [201-]). Os critérios de seleção de bibliotecas foram definidos de acordo com os objetivos iniciais do estudo, sendo eles, ser originada de células ou tecidos renais de humanos ou camundongos após exposição a substância nefrotóxica, os tecidos deveriam ser originados de animais saudáveis para evitar a inserção de variantes advindas de condições como, por exemplo, doenças. Durante a busca, foram excluídas metodologias indesejáveis como sequenciamento de miRNA e imunoprecipitação de cromatina, por exemplo.

A ferramenta Entrez Direct (Edirect), versão 7.00, que permite acessar informações das bases de dados do NCBI a partir de um terminal UNIX, foi utilizada para a busca (KANS, 2010). A *query* utilizada foi composta por uma lista de fármacos sabidamente nefrotóxicos, termos relacionados a rins e toxicidade, limitando-se às espécies supracitadas, além de excluir metodologias de sequenciamento indesejadas. A lista completa dos fármacos utilizados na busca pode ser encontrada no Apêndice A. Abaixo, os comandos utilizados na busca:

```
$ esearch -db sra -query "(fármacos separados pelo operador booleano OR) AND (kidney OR renal OR nephron OR nephrotoxicity) AND (Homo sapiens[ORGN] OR Mus musculus[ORGN]) NOT (miRNA OR ChIP-Seq OR Ribosome profiling OR siRNA OR CRISPR)" | efetch -format uid
```

A fim de obter o maior número de dados possível e garantir a qualidade e abrangência do trabalho, novas buscas no banco de dados SRA foram realizadas periodicamente, limitando-se à espécie e fármaco identificados inicialmente, a saber, *Mus musculus* (camundongo) e cisplatina, respectivamente, e mantendo os demais critérios da pesquisa inicial.

O *download* das bibliotecas foi realizado utilizando o *software* Aspera Connect, versão 3.6.2.117442, ascp, versão 3.5.6.112424 (NCBI, 2014), e a conversão para o formato FASTQ foi realizado com a ferramenta fastq-dump, versões 2.5.5 e 2.9.2, integrante do pacote de ferramentas SRA Toolkit (NCBI, [201-]), incluindo a opção `--split-files` nos casos de sequenciamento de *reads* pareadas (*paired-end*). O controle de qualidade dos sequenciamentos foi realizado com o programa FastQC (BABRAHAM BIOINFORMATICS), versão v0.11.4, e a visualização dos resultados foi realizada com o programa MultiQC (EWELS et al., 2016), versão 1.8.dev0, que permite agregar resultados de diversas amostras em imagens únicas.

3.3 Alinhamento de *reads* contra o genoma de referência

O alinhamento de *reads* foi realizado com o programa STAR – *Spliced Transcripts Alignment to a Reference* (DOBIN et al., 2012), versão 2.5.3a, que utiliza a busca por semente (*seed searching step*) e agrupamento e pontuação (*clustering/stitching/scoring step*) como os dois principais passos para realização do alinhamento. Resumidamente, o primeiro passo consiste em encontrar a maior sequência mapeável possível para cada *read* em relação à posição genômica. Por se tratar de alinhamento de *reads* de RNA, ao encontrar uma região intrônica do genoma essa região não terá correspondência na *read* que está sendo alinhada. Assim, a busca pela maior sequência mapeável continua por esta região até que o alinhamento volte a ser possível, ou seja, até que seja encontrada a próxima junção de *splicing* (DOBIN et al., 2012).

O segundo passo agrupa e une as sequências alinhadas no passo anterior por proximidade genômica (janelas genômicas), sendo que o algoritmo que une as sequências permite qualquer número de incompatibilidade (*mismatch*), mas apenas uma inserção ou deleção. A união das sequências alinhadas é determinada por pontuação de alinhamento local, sendo que a união que obtiver maior pontuação é a escolhida como melhor alinhamento. Em comparação a outros programas disponíveis, o STAR possui maior velocidade de alinhamento, mantendo a precisão e sensibilidade elevadas (DOBIN et al., 2012).

No presente trabalho, o alinhamento foi realizado contra o genoma de referência do camundongo GRCm38.p6. Considerando que a precisão do mapeamento é aumentada quando fornecidas as informações de junção éxon-íntron para o STAR, utilizou-se a anotação do genoma, em formato GTF, versão M20 disponibilizada pelo GENCODE. Apenas *reads* mapeadas unicamente foram incluídas no resultado final do STAR.

3.4 Pseudoalinhamento contra o transcriptoma de referência e estimativa de abundância de transcritos

Tendo em vista o objetivo de analisar a expressão gênica ao nível de transcritos, optou-se por utilizar um programa que realizasse a quantificação de transcritos em detrimento de programas que quantificam ao nível de gene. Assim, foi utilizado o programa kallisto (BRAY et al., 2016), versão 0.43.1. Este programa utiliza um algoritmo de pseudoalinhamento para realizar a quantificação dos transcritos, no qual um grafo de Brujin

do transcriptoma é construído, sendo os nós correspondentes aos *k-mers* do transcriptoma e os transcritos representados pelos caminhos do grafo que conectam os *k-mers* que o compõem. Para atribuir uma *read* a um transcrito, o kallisto ignora *k-mers* redundantes, o que leva à diminuição do risco de atribuições incorretas. Em comparação a outros programas disponíveis, o kallisto apresenta exatidão tão boa quanto ou melhor na quantificação de transcritos, mesmo nos genes que apresentam muitas isoformas. Além disso, apresenta alta velocidade de execução com baixo uso de memória (BRAY et al., 2016; ZHANG et al., 2017).

O índice do kallisto foi gerado a partir do transcriptoma do camundongo disponibilizado pelo GENCODE, versão M20, em formato FASTA. A função “quant” do kallisto foi utilizada para a realização do pseudoalinhamento e obtenção da abundância dos transcritos. O kallisto utiliza diretamente os arquivos em formato FASTQ e seu modo *default* considera o sequenciamento de *reads* pareadas (*paired-end*). Como foram obtidas tanto bibliotecas com *reads* pareadas como não pareadas (*single-end*), nos casos onde as bibliotecas eram de *reads* não pareadas, foram incluídos os parâmetros --single, -l e -s, onde o primeiro parâmetro indica que as *reads* não são pareadas, o segundo refere-se ao tamanho médio dos fragmentos utilizados na construção das bibliotecas e o último parâmetro refere-se ao desvio padrão do tamanho dos fragmentos. Informações sobre o tamanho dos fragmentos e seu desvio padrão não estavam disponíveis, dessa maneira optou-se por utilizar os valores de 200 pb e 20 pb, respectivamente. A escolha desses valores se deu tanto devido à indicação do manual do programa kallisto (PACHTER LAB, [201-]) quanto pelas informações disponíveis no guia de preparo de amostras referente ao kit utilizado pelos autores (ILLUMINA, 2012; GALGAMUWA et al., 2016). Em todos os casos, *reads* pareadas e não pareadas, o parâmetro -b foi utilizado com o valor 100 para que cem reamostragens com reposição fossem realizadas no passo de quantificação, permitindo mensurar a incerteza das estimativas de abundância dos transcritos.

3.5 Expressão diferencial de transcritos

A análise de expressão diferencial foi realizada com o sleuth (PIMENTEL et al., 2017), versão 0.30.0, um pacote em R desenvolvido para analisar experimentos de RNA-seq onde a abundância foi quantificada utilizando-se o kallisto. O sleuth considera que a variabilidade contida numa amostra é decorrente de dois componentes aditivos, sendo eles a

variância biológica e a variância inferencial. A variância biológica é aquela decorrente das diferenças de expressão entre as amostras e da preparação das bibliotecas. Já a variância inferencial é derivada dos procedimentos computacionais e do sequenciamento aleatório de fragmentos. Utilizando os resultados gerados pelo kallisto através das reamostragens (*bootstraps*), o sleuth é capaz de desacoplar a variância biológica da inferencial. Com isso, a incerteza das quantificações é levada em consideração na análise de expressão diferencial, tornando o sleuth mais exato, tanto ao nível de gene quanto de transcrito, que outras abordagens disponíveis (PIMENTEL et al., 2017).

A identificação de transcritos diferencialmente expressos (TDE) foi realizada a partir do teste de razão de verossimilhança (TRV; *sleuth_lrt()*) e do teste de Wald (*sleuth_wt*). O TRV compara dois modelos previamente gerados, sendo eles o modelo completo e o modelo reduzido. O modelo completo é aquele que considera a condição de interesse, controle *versus* tratamento com cisplatina, no caso deste trabalho, enquanto o modelo reduzido é a hipótese nula. O valor-p gerado para cada transcrito na comparação entre os modelos é ajustado para múltiplas comparações gerando o valor-q (*qvalue*), que foi utilizado como corte de significância para identificação dos TDEs. No entanto, o TRV não calcula o valor de *fold change*, o que inviabiliza a comparação de expressão entre as condições. Para tanto, foi utilizado o teste de Wald. Numa análise de desempenho, Hauck e Donner (1977) (HAUCK; DONNER, 1977) observaram que o teste de Wald pode não rejeitar a hipótese nula quando o coeficiente é significativo, sendo o TRV recomendado para análise de significância. Assim, foram considerados diferencialmente expressos aqueles transcritos com valor-q menor que 0,01 e *fold change* maior que 2 para superexpressão ($\log_2 \text{fold change} > 1$) e menor que -2 para subexpressão ($\log_2 \text{fold change} < -1$).

3.6 Identificação de transcrição *readthrough*

A transcrição *readthrough*, ou seja, a transcrição contínua pela região intergênica devido ao não término da transcrição pela Pol II, foi acessada utilizando-se o programa DoGFinder (WIESEL; SABATH; SHALGI, 2018), versão 1.0.0, que identifica e quantifica esse tipo de transcrição a partir de dados gerados por sequenciamento de RNA. Para tanto, o programa utiliza arquivos de alinhamento em formato BAM e a anotação do genoma em formato GTF.

Um dos passos necessários para identificar a transcrição em regiões que não contêm genes é conhecer quais regiões contêm genes. O programa DogFinder obtém essa informação através de seu *script Get_loci_annotation* utilizando anotações do genoma em formato GTF obtidas no *site* da UCSC (*University of California, Santa Cruz* - <https://genome.ucsc.edu/cgi-bin/hgTables>). Para este trabalho optou-se por duas anotações, ambas pertencentes ao grupo de genes e genes preditos tendo como referência a montagem GRCm38/mm10 do genoma do camundongo, sendo uma a tabela *Basic* do All GENCODE VM18 e a outra a tabela UCSC RefSeq (refGene) do NCBI RefSeq. Assim, o primeiro passo realizado no programa DogFinder foi construir uma anotação única em formato BED a partir das duas anotações fornecidas. Genes sobrepostos nas duas anotações são fundidos em um único locus e, caso os limites do gene sejam distintos nas anotações, o programa utiliza os limites que abrangem a maior região.

Na sequência, os arquivos de alinhamento em formato BAM gerados pelo programa STAR passaram pela etapa de pré-processamento (*Pre_process*). Para que os níveis de expressão dos DoGs possam ser comparados, os arquivos BAM passam por uma amostragem que limita o número de *reads* utilizadas nas próximas etapas. Além disso, a profundidade do sequenciamento pode influenciar na identificação dos DoGs, uma vez que um dos critérios utilizados é a cobertura contínua.

A identificação dos DoGs, realizada com o *script Get_DoGs*, utiliza o arquivo de anotação em formato BED gerado no primeiro passo para excluir os loci gênicos dos arquivos BAM pré-processados gerados no segundo passo, gerando um arquivo BED com as informações apenas das regiões não gênicas. Este arquivo foi então utilizado para a identificação dos DoGs candidatos, seguindo os parâmetros *-minDoGLen 4000*, *-minDoGCov 0.8*, que se referem, respectivamente, ao tamanho mínimo de 4000 pb contados a partir da extremidade 3' do gene anterior e 80% de cobertura. O alongamento dos candidatos identificados foi configurado com os parâmetros *-w 200* para que a extensão fosse realizada com uma janela deslizante de 200 pb, com sobreposição, e *-mode F* para que a cobertura mínima de 80% fosse computada sobre o comprimento total do DoG. Além do parâmetro *-s*, que indica a existência de especificidade de fita das bibliotecas.

O passo acima descrito gerou arquivos de anotações de DoGs em formato BED para cada uma das amostras. Esses arquivos foram utilizados para gerar uma anotação única contendo todos os DoGs identificados nas amostras (*script Union_DoGs_annotation*). A anotação unificada foi utilizada no último passo (*script Get_DoGs_rpkms*) para calcular o

nível de expressão dos DoGs a partir dos arquivos BAM pré-processados. A medida adotada pelo DoGFinder é o RPKM (*reads* por quilobase por milhão de *reads* mapeadas).

A fim de comparar o número de DoGs encontrados especificamente em cada uma das condições, o *script Common_DoGs_annotation* foi utilizado. Aqui, o programa identifica quais DoGs são comuns às anotações de DoG utilizadas como entrada. Portanto, essa etapa foi executada três vezes, sendo uma com as anotações de DoGs referentes às amostras do grupo controle, outra com as anotações referentes às amostras tratadas e, por fim, com as anotações referentes a todas as amostras.

Devido às características do programa e dos estudos utilizados neste trabalho, optou-se por utilizar apenas bibliotecas com especificidade de fitas para a caracterização da transcrição *readthrough*. Uma vez que o programa retira dos arquivos BAM todos os loci gênicos disponibilizados nas anotações, a identificação dos DoGs torna-se mais precisa com a utilização de bibliotecas com especificidade de fita.

3.7 Análise funcional

A análise de função dos transcritos diferencialmente expressos foi dividida em dois processos devido às peculiaridades de genes e transcritos codificadores e não codificadores de proteínas. Em suma, os transcritos codificadores de proteínas foram analisados com o enriquecimento de processos biológicos descritos pelo *Gene Ontology* (GO). Já os transcritos provenientes de genes não codificadores de proteínas tiveram suas funções analisadas a partir de informações disponíveis na literatura. Os genes que sofreram transcrição *readthrough* também foram analisados pelo enriquecimento de processos biológicos do *Gene Ontology*.

3.7.1 Genes codificadores de proteínas

A ontologia gênica utilizada neste trabalho foi aquela definida pelo *Gene Ontology Consortium* para processos biológicos. Entende-se por processo biológico o objetivo biológico ao qual o gene ou produto gênico contribui, podendo envolver transformações químicas ou físicas (ASHBURNER et al., 2000; THE GENE ONTOLOGY CONSORTIUM, 2019).

O pacote R topGO (ALEXA; RAHNENFÜHRER, 2019), versão 2.36.0, foi utilizado para fazer o enriquecimento de processos biológicos presentes nos TDE. Para identificar os

processos biológicos enriquecidos, os TDE foram divididos em quatro subgrupos. Inicialmente, foram separados entre superexpressos e subexpressos e cada um desses grandes grupos foi subdividido em outros dois: um contendo os transcritos codificadores de proteínas e outro contendo os transcritos originados de genes codificadores de proteínas e que não são classificados como codificadores de proteínas como, por exemplo, transcritos classificados como íntron retido, NMD e transcrito processado. Cada um dos quatro grupos foi analisado individualmente, mantendo as mesmas configurações para todos. Importante ressaltar que os TDE utilizados no enriquecimento de processos biológicos foram aqueles comuns aos estudos identificados como descrito no item “3.2 Aquisição de dados de sequenciamento de RNA e controle de qualidade”.

Antes de realizar a análise de enriquecimento propriamente dita, é necessário construir um objeto *topGOdata* (*new("topGOdata", ...)*). Esse objeto deve conter todas as informações que serão utilizadas na etapa de enriquecimento. Assim, a construção do mesmo foi realizada utilizando uma lista nomeada de genes, o pacote de anotação *org.Mm.eg.db*, versão 3.8.2, para designação dos termos GO para cada gene (CARLSON, 2019) e considerando apenas nós (termos GO) com um mínimo de dez genes.

A lista de genes deve ser composta por aqueles de interesse e por outros que não são de interesse. Os genes de interesse são aqueles diferencialmente expressos e os que não são de interesse são chamados de *background*. O *background* foi composto por transcritos não diferencialmente expressos que apresentaram média de contagem similar aos TDE. Dessa forma, o *background* foi composto apenas por transcritos que foram expressos nas amostras. Apenas transcritos ou genes classificados como codificadores de proteínas foram utilizados na etapa de construção do *background*. Além disso, os *backgrounds* foram primeiro construídos para cada um dos estudos e, em seguida, foram unificados removendo transcritos duplicados. Por fim, os identificadores de transcritos Ensembl foram substituídos por seus respectivos identificadores de genes Ensembl, para que pudessem ser utilizados na etapa subsequente de enriquecimento.

topGO disponibiliza diversos algoritmos para realização do enriquecimento de termos, dentre eles estão os algoritmos *elim* e *weight*. Ambos algoritmos consideram a topologia dos termos GO no cálculo da significância. O método *elim* é um método de eliminação, onde os genes presentes em nós mais baixos, ou seja, mais específicos são removidos de todos os nós ancestrais (mais gerais). Já o método *weight*, compara os valores de significância entre termos GO conectados, quando a significância é maior para o nó mais específico, o peso dos genes que compõe esse nó é removido do cálculo de significância dos nós mais gerais conectados a

ele. Na prática, a diferença entre os dois métodos é que o primeiro apresenta menor taxa de falsos positivos que o segundo, em contrapartida, possui maior risco de descartar termos verdadeiramente enriquecidos (ALEXA; RAHNENFÜHRER; LENGAUER, 2006). Na etapa de enriquecimento (*runTest()*), foi utilizado o teste exato de Fisher para detecção dos termos enriquecidos com o algoritmo *weight01*, que é o método padrão do topGO, sendo este método uma mistura dos métodos *elim* e *weight*.

Os resultados foram filtrados para os cinquenta nós mais significativos de acordo com o valor-p, sendo 0,01 o menor valor aceito. Para efeito de visualização, os processos biológicos enriquecidos foram curados manualmente, retirando-se as redundâncias.

3.7.2 Genes não codificadores de proteínas

Para acessar funções conhecidas de TDE provenientes de genes não codificadores, utilizou-se as informações disponíveis no banco de dados NONCODEv5. NONCODEv5 é um banco de dados de lncRNA que armazena dados de dezessete espécies, incluindo camundongo. Atualmente, esse banco de dados possui mais de oitenta mil genes e mais de cento e trinta mil transcritos não codificadores de camundongos com seu respectivo perfil de expressão tecidual e, para alguns desses genes, também é possível encontrar as funções preditas (FANG et al., 2018). Para identificar as funções dos TDE provenientes de genes não codificadores, foi feito o *download* da tabela “*Predicted Functions of NONCODE*” disponível em <<http://www.noncode.org/download.php>>. Como o NONCODE utiliza identificadores próprios para os lncRNAs, realizou-se a conversão dos identificadores Ensembl dos genes para identificadores NONCODE através da função *Conversion* disponível em <http://www.noncode.org/id_conversion.php>. Em seguida, foi realizada a busca pelos identificadores NONCODE dos TDE na tabela de função predita.

Além da busca pelas informações contidas no banco de dados NONCODE, os transcritos originados de genes não codificadores de proteínas tiveram suas funções analisadas a partir de revisão bibliográfica. Para tanto, realizou-se buscas por publicações indexadas na base de dados PubMed com *queries* compostas pelos identificadores Ensembl e NONCODE para gene e transcrito, nome do gene e nome do transcrito separados pelo operador booleano *OR*. Por exemplo, para o lincRNA Trp53cor1, utilizou-se a *query* a seguir:

```
ENSMUST00000133221 OR ENSMUSG00000085912 OR NONMMUT028986 OR NONMMUG017913
OR Trp53cor1-201 OR Trp53cor1 OR Gm16197 OR lincRNA-p21
```

3.7.3 Genes que sofreram transcrição *readthrough*

O enriquecimento de processos biológicos dos genes que sofreram transcrição *readthrough* foi realizado utilizando a metodologia descrita no item “3.7.1 Genes codificadores de proteínas”, à exceção da construção do *background*. Nesse caso, o *background* foi composto por transcritos não sofreram transcrição *readthrough* que apresentaram média de contagem similar àqueles que sofreram. Como a análise de transcrição *readthrough* foi realizada apenas para as bibliotecas com especificidade de fita, o *background* foi construído apenas com dados advindos dessas bibliotecas. Foram considerados significativos os processos biológicos com valor-p menor que 0,01.

3.8 Visualização dos resultados

Os gráficos apresentados na sessão “4. Resultados” foram gerados utilizando pacotes do R, conforme abaixo:

- Diagramas de Venn: VennDiagram, versão 1.6.20
- Gráficos de barras: ggplot2, versão 3.2.1
- Gráficos tipo vulcão: ggplot2, versão 3.2.1
- *Heatmaps*: pheatmap, versão 1.0.12
- PCA: sleuth, função plot_pca(), versão 0.30.0
- Upsets: upSetR, versão 1.4.0

3.9 Recursos computacionais

Os dados foram processados em servidor local e no *cluster* HPC sagarana, CEPAD-ICB-UFMG. A paralelização, quando necessária, foi realizada com o software GNU Parallel (TANGE, 2011).

4 RESULTADOS

4.1 Caracterização dos estudos analisados

As buscas realizadas no SRA retornaram dois estudos de interesse, ambos depositados no *Gene Expression Omnibus* (GEO), base de dados mantida pelo NCBI, e podem ser acessados através dos números de acesso GSE69652 e GSE106993. Os artigos “*Dichloroacetate prevents cisplatin-induced nephrotoxicity without compromising cisplatin anticancer properties*” (GALGAMUWA et al., 2016) e “*The proteome microenvironment determines the protective effect of preconditioning in cisplatin-induced acute kidney injury*” (SPÄTH et al., 2019) foram publicados a partir das análises dos dados disponíveis em GSE69652 e GSE106993, respectivamente, pelos grupos que geraram os dados.

Por questões de clareza, no presente trabalho denominou-se “Estudo 1” as menções aos dados disponíveis em GSE69652 e “Estudo 2” as menções aos dados disponíveis em GSE106993. Na Tabela1, pode-se ver características dos desenhos experimentais de ambos. Os desenhos experimentais seguem as necessidades dos estudos originais e, portanto, diferem em diversos aspectos como, por exemplo, linhagem de camundongo e tempo de tratamento. Já a espécie, medicação utilizada e dose aplicada nos animais são comuns em ambos estudos.

Tabela 1. Características do desenho experimental dos estudos analisados

Características	Estudo 1 GSE69652	Estudo 2 GSE106993
Linhagem dos camundongos	BALB/c	C57BL/6
Gênero	Machos	Machos
Idade (semanas)	6-10	19-21
Número de animais por grupo	3	4
Dose de cisplatina (mg/kg)	20	20
Via de administração	Intraperitoneal	Intraperitoneal
Tempo de tratamento	24 horas	72 horas
Seleção de RNA	Cauda poliA	Depleção de rRNA
Plataforma de sequenciamento	HiSeq2500	HiSeq4000
Estratégia de sequenciamento	<i>Single-end</i>	<i>Paired-end</i>
Especificidade de fita	Não específico	Específico
Tamanho das reads	101 pb	2x76 pb
Reads do grupo controle (nº médio)	49.495.926	85.472.516
Reads do grupo tratado (nº médio)	33.048.786	82.509.418

Os dois estudos tiveram como objetivo principal avaliar pré-tratamento ou condicionamentos que pudessem mitigar a nefrotoxicidade induzida pela cisplatina. Portanto, foram sequenciados RNAs de rins de animais tratados com cisplatina após pré-tratamento ou após submissão aos condicionamentos do estudo, além dos RNAs dos rins de animais que não foram submetidos a nenhum condicionamento. Para atender aos objetivos deste trabalho, apenas amostras de rins de animais tratados com cisplatina ou salina (controle) que não tenham sido submetidos a algum pré-tratamento ou condicionamento foram analisadas. As amostras utilizadas podem ser vistas na Tabela 2, onde também são encontradas as denominações das amostras (coluna Amostra) utilizadas neste trabalho.

Tabela 2. Identificação das amostras

Estudo	Acesso GEO	Corrida SRA	Condição	Amostra
Estudo 1	GSE69652	SRR2056115	Controle	Controle A
		SRR2056116		Controle B
		SRR2056117		Controle C
		SRR2056121	Tratado	Cisplatina A
		SRR2056122		Cisplatina B
		SRR2056123		Cisplatina C
Estudo 2	GSE106993	SRR6297964	Controle	Controle D
		SRR6297965		Controle E
		SRR6297966		Controle F
		SRR6297967		Controle G
		SRR6297956	Tratado	Cisplatina D
		SRR6297957		Cisplatina E
		SRR6297958		Cisplatina F
		SRR6297959		Cisplatina G

Todas as amostras foram analisadas quanto à qualidade com o programa FastQC apresentando excelente qualidade de sequenciamento (Figura 12).

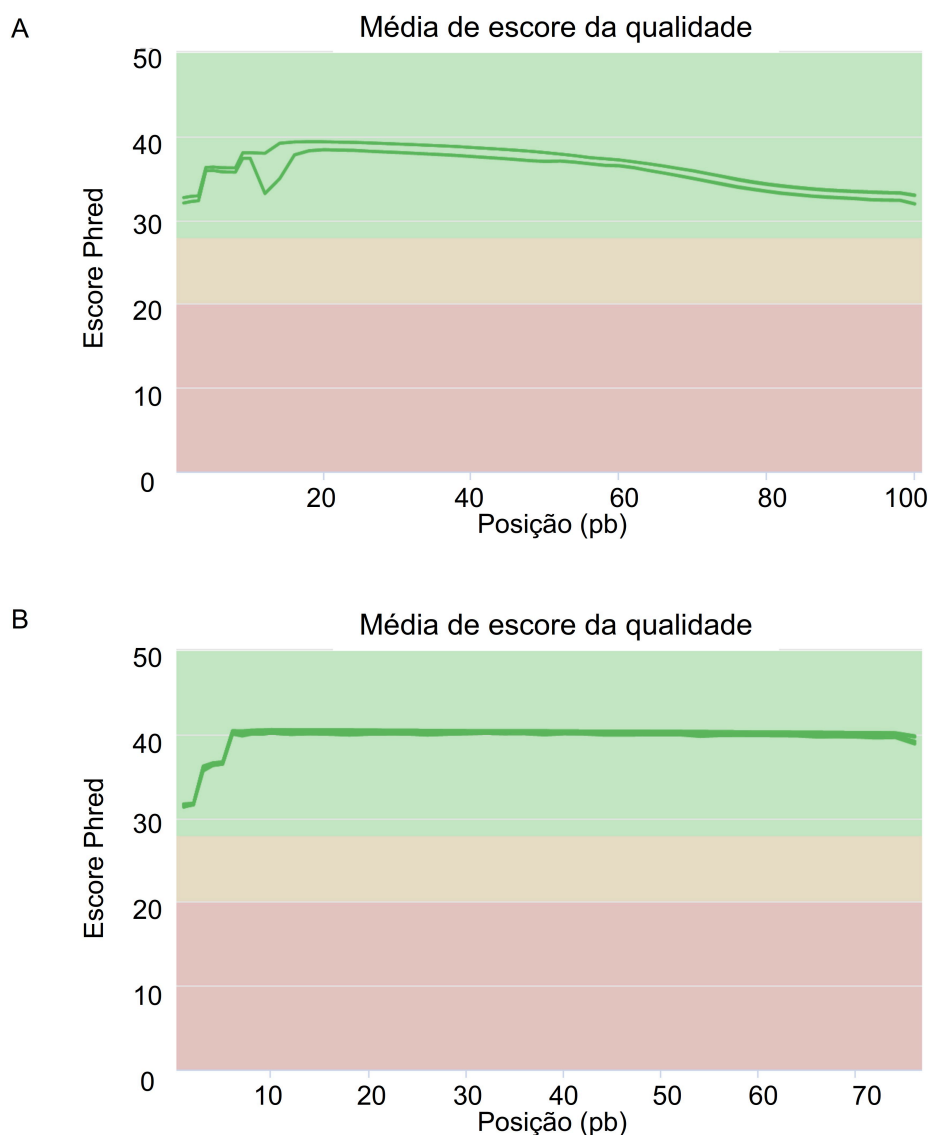


Figura 12. Qualidade dos sequenciamentos. (A) Média de qualidade por base das amostras do Estudo 1. (B) Média de qualidade por base das amostras do Estudo 2. Linhas em verde representam cada uma das amostras dos Estudos. pb, pares de bases.

4.2 Caracterização das amostras

Todas as amostras foram submetidas ao alinhamento e pseudoalinhamento de *reads* independentemente. Tanto no alinhamento realizado com o programa STAR quanto no pseudoalinhamento gerado pelo programa kallisto, os percentuais de mapeamento entre as amostras controle e as amostras tratadas foi similar no Estudo 1, sendo um pouco maior nas amostras do grupo controle. Já no Estudo 2, essa diferença percentual entre os grupos foi mais marcante (Tabela 3).

Tabela 3. Percentual de mapeamento de reads

Estudo	Amostra	Total de <i>reads</i> (n)	<i>Reads</i> mapeadas (n)		Mapeamento (%)	
			STAR	kallisto	STAR	kallisto
Estudo 1	Controle A	48.811.345	42.632.250	43.510.848	87,34	89,14
	Controle B	58.518.810	50.724.080	51.852.051	86,68	88,61
	Controle C	41.157.623	35.965.169	36.695.688	87,38	89,16
	Cisplatina A	31.803.541	27.811.884	28.069.330	87,45	88,26
	Cisplatina B	30.463.903	26.735.824	27.018.200	87,76	88,69
	Cisplatina C	36.878.915	32.375.801	32.792.475	87,79	88,92
Estudo 2	Controle D	87.176.036	76.308.581	65.897.519	87,53	75,59
	Controle E	85.377.925	75.106.837	65.496.437	87,97	76,71
	Controle F	89.918.107	78.439.944	67.593.673	87,23	75,17
	Controle G	79.417.997	69.773.490	60.739.018	87,86	76,48
	Cisplatina D	78.826.148	65.556.463	56.620.198	83,17	71,83
	Cisplatina E	76.793.102	62.488.298	55.204.753	81,37	71,89
	Cisplatina F	80.634.130	66.843.820	58.318.105	82,90	72,32
Cisplatina G	93.784.295	79.287.722	66.151.458	84,54	70,54	

Para realizar as análises de expressão diferencial, o programa sleuth realiza uma normalização das estimativas de abundância obtidas pelo programa kallisto de forma que esses valores se tornem comparáveis entre as amostras analisadas conjuntamente. A partir desses valores normalizados, foram gerados gráficos de análise de componentes principais (PCA), nos quais é possível visualizar o comportamento das amostras frente ao tratamento (Figura 13). As amostras dos Estudos 1 e 2 foram submetidas à normalização das estimativas de abundância individualmente e também em conjunto. Na análise em conjunto, as diferenças entre os estudos foram incluídas no modelo nulo para que, ao acessar as diferenças de expressão, estas refletissem a diferença entre a presença e ausência de tratamento com cisplatina. Ainda assim, devido às diferenças entre os estudos, a componente principal 1 separou as amostras entre Estudo 1 e Estudo 2 (Figura 13A), enquanto a componente principal 2 separou as amostras pelo tratamento (Figura 13B). A análise de componentes principais individual dos estudos mostra que a componente principal 1 separa adequadamente as amostras pelo tratamento (Figura 13C e 13E). Os valores normalizados de estimativa de abundância também foram utilizados para construir *heatmaps* de distância euclidiana das

amostras. Tanto para o Estudo 1 quanto para o Estudo 2, os dendrogramas mostram que as amostras se agrupam adequadamente pelo tipo de tratamento (Figura 13D e 13F).

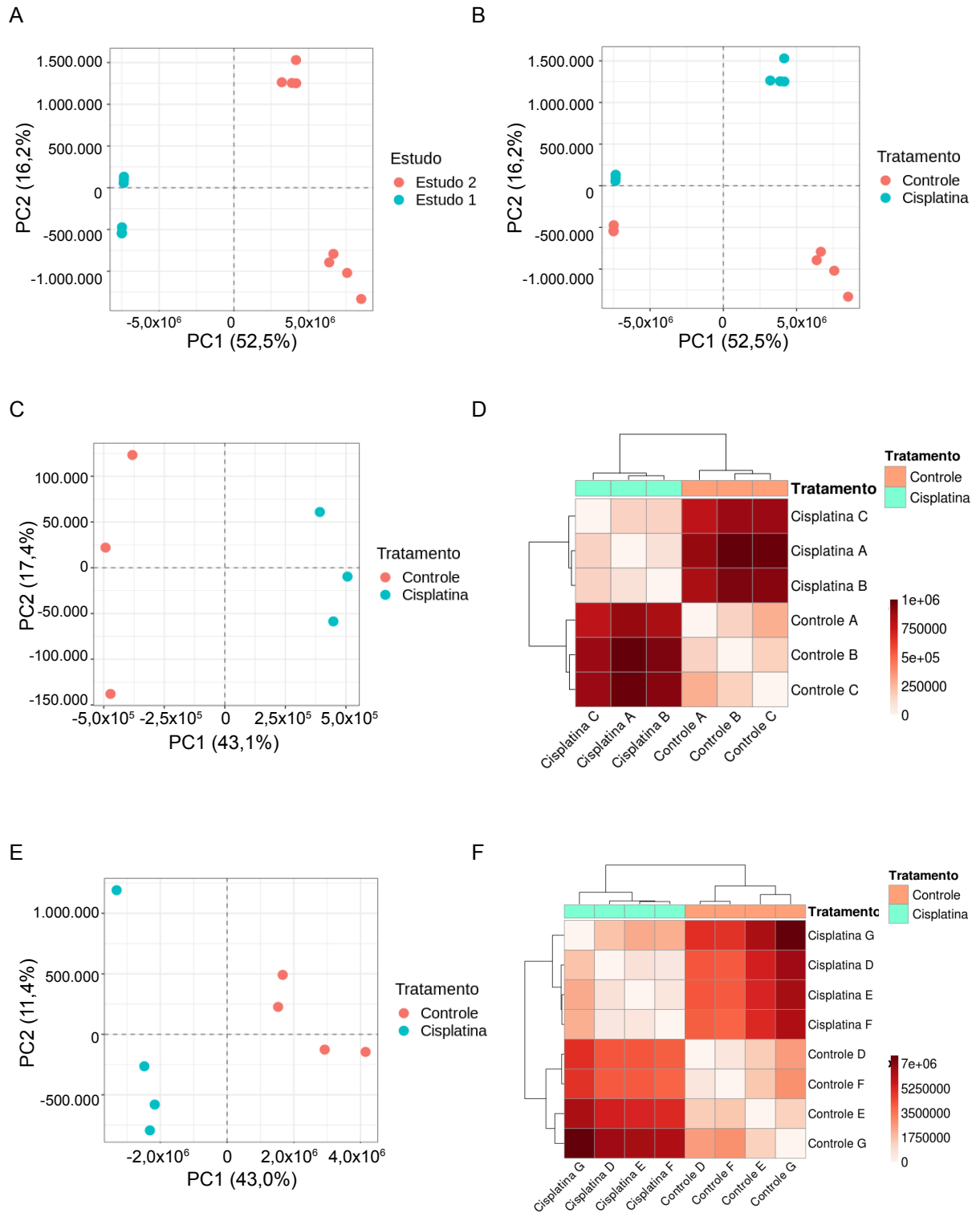


Figura 13. Perfil de distribuição de amostras. (A-B) PCA de análise conjunta dos Estudos 1 e 2. Em A, as amostras foram coloridas de acordo com o estudo a que pertencem, destacando que PC1 separa as amostras pelo estudo. Em B, as amostras foram coloridas de acordo com o tratamento, independente do estudo, destacando que PC2 separa as amostras pelo tratamento. (C-D) PCA e *heatmap* de amostras do Estudo 1. (E-F) PCA e *heatmap* de amostras do Estudo 2. Escala de cor dos *heatmaps* representa a distância euclidiana das amostras, onde quanto menor o valor maior a similaridade das amostras. PCA, análise de componentes principais. PC1, componente principal 1. PC2, componente principal 2.

Devido à impossibilidade de agrupar adequadamente as amostras pelo tratamento quando os estudos são analisados conjuntamente, optou-se por seguir o trabalho analisando os mesmos individualmente.

4.3 Estudo do transcriptoma renal após tratamento com cisplatina

4.3.1 Perfil de expressão diferencial de transcritos

A expressão diferencial de transcritos foi acessada através do teste de Wald implementado pelo programa sleuth. Observou-se que o Estudo 2 apresentou maior número de TDE que o Estudo 1 e que, em ambos, os TDE subexpressos superam os superexpressos. O perfil geral de expressão diferencial pode ser visto na Figura 14.

Para as demais análises, os transcritos foram separados em três grandes grupos de acordo com a classificação de transcritos utilizada pelo Ensembl. O grupo de transcritos codificadores de proteínas (CP) é composto pelos transcritos classificados como codificadores de proteínas. Já o grupo de transcritos processados originados de genes codificadores de proteínas (TPCP) contém os transcritos que, apesar de serem oriundos de genes codificadores de proteínas, não codificam proteínas ou que codificam apenas pequenos polipeptídeos e que seguem para via de degradação NMD. Por fim, o grupo de transcritos não codificadores de proteínas (NCP) é composto pelos transcritos não codificadores gerados a partir de genes que não são classificados como codificadores de proteínas.

O Estudo 1 apresentou 722 transcritos superexpressos e 1063 transcritos subexpressos, enquanto o Estudo 2 apresentou 3253 transcritos superexpressos e 3379 transcritos subexpressos. O número de transcritos subexpressos supera o de superexpressos em todos os grupos de transcritos, exceto no grupo TPCP do Estudo 2. A distribuição dos transcritos pelo *fold change* de cada um dos Estudos está representada em gráficos tipo vulcão (Figura 15).

Os dois estudos apresentaram, em comum, 404 transcritos superexpressos, sendo 288 CP, 77 TPCP e 39 NCP, e 637 transcritos subexpressos, sendo 445 CP, 126 TPCP e 66 NCP (Figura 16, Apêndices B e C).

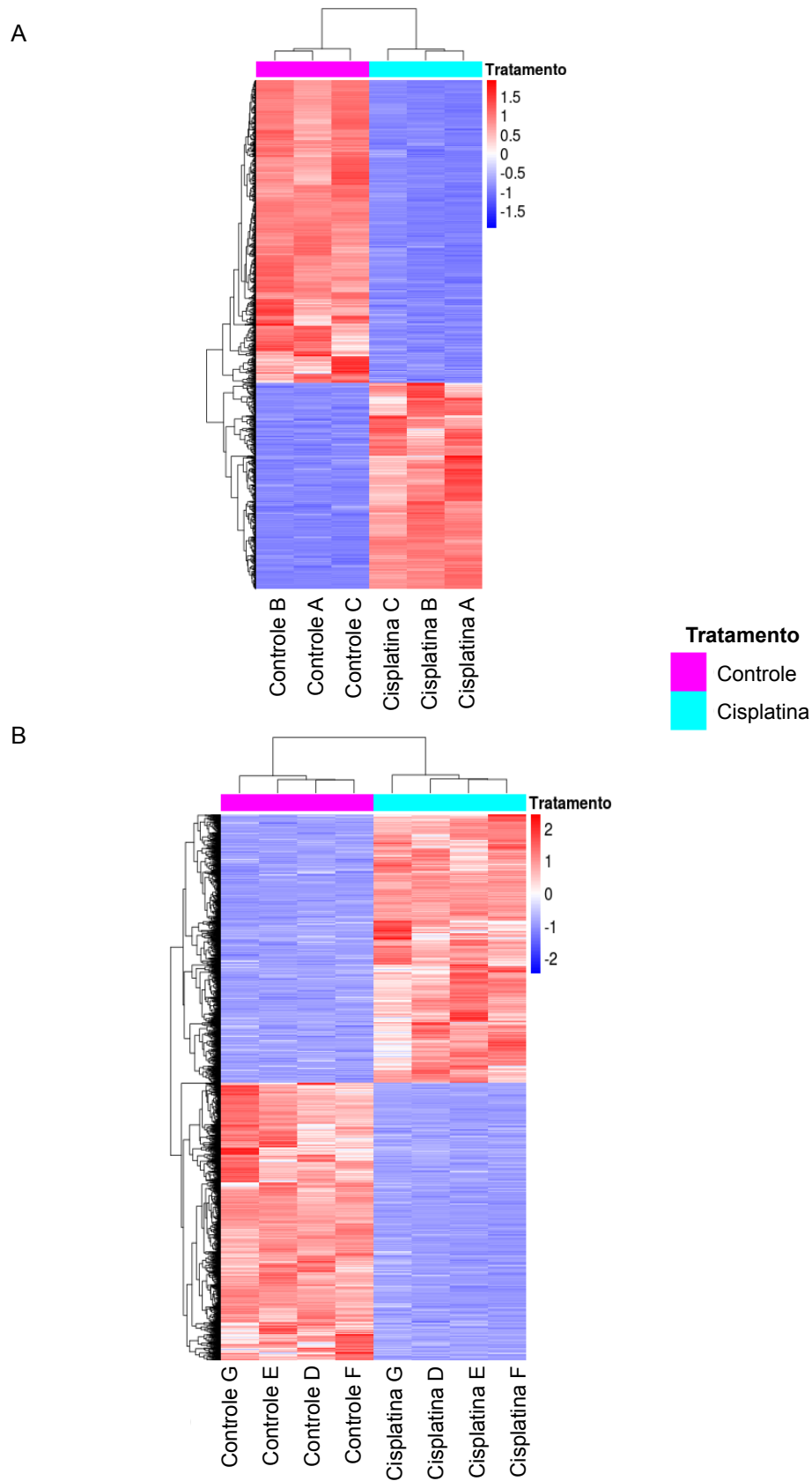


Figura 14. Heatmaps de transcritos diferencialmente expressos. (A) Transcritos diferencialmente expressos no Estudo 1, sendo 722 superexpressos e 1063 subexpressos. (B) Transcritos diferencialmente expressos no Estudo 2, sendo 3253 superexpressos e 3379 subexpressos. Escala de cor representa o *z-score* das estimativas de abundância dos transcritos.

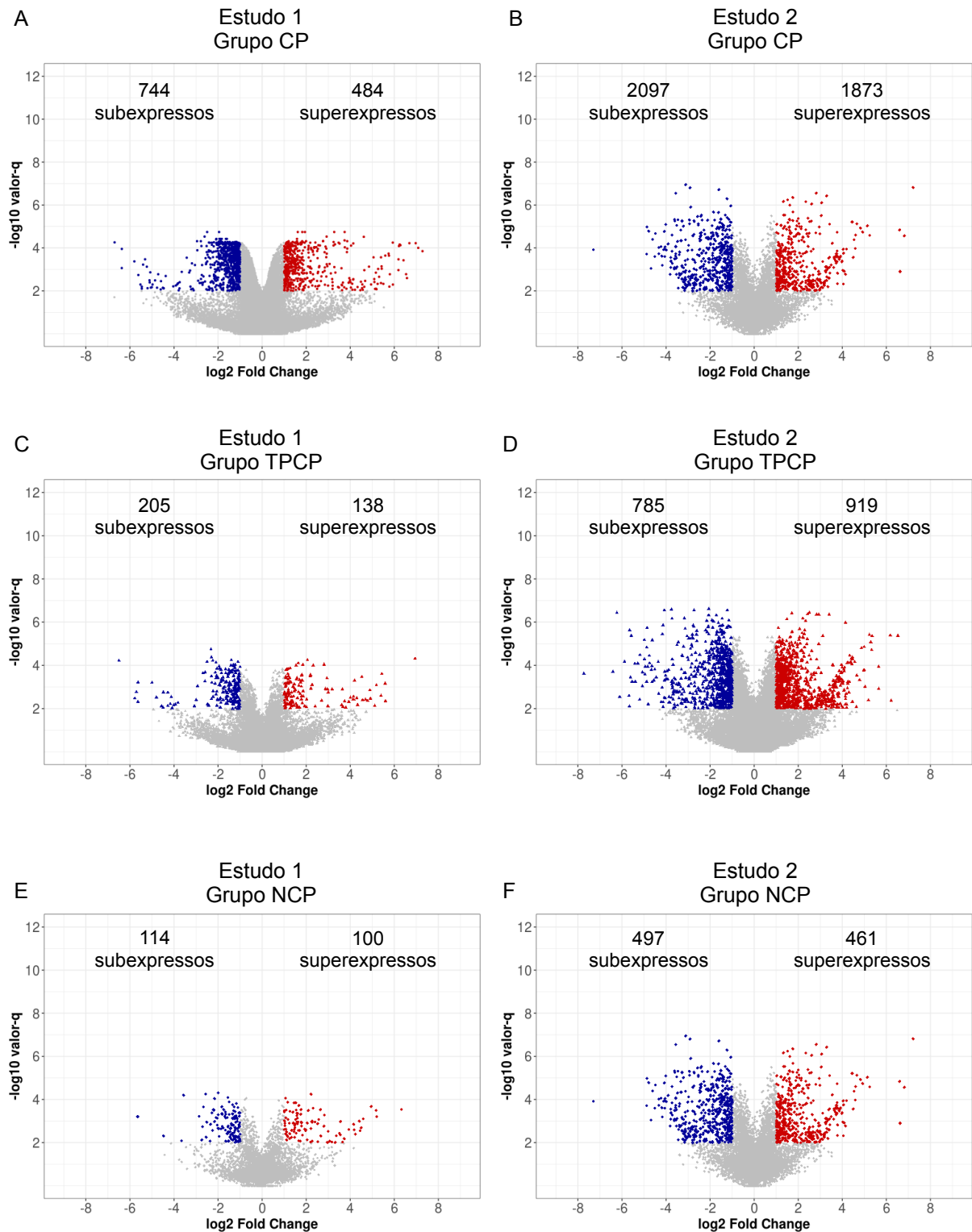


Figura 15. Dispersão de *fold change* dos transcritos diferencialmente expressos. (A) TDE codificadores de proteínas do Estudo 1. (B) TDE codificadores de proteínas do Estudo 2. (C) TDE não codificadores originados de genes codificadores de proteínas do Estudo 1. (D) TDE não codificadores originados de genes codificadores de proteínas do Estudo 2. (E) TDE não codificadores do Estudo 1. (F) TDE não codificadores do Estudo 2. CP, codificadores de proteína. NCP, não codificadores de proteína. TDE, transcritos diferencialmente expressos. TPCP, transcritos processados originados de genes codificadores de proteínas. Pontos vermelhos, azuis e cinzas representam transcritos superexpressos, subexpressos e não diferencialmente expressos, respectivamente.

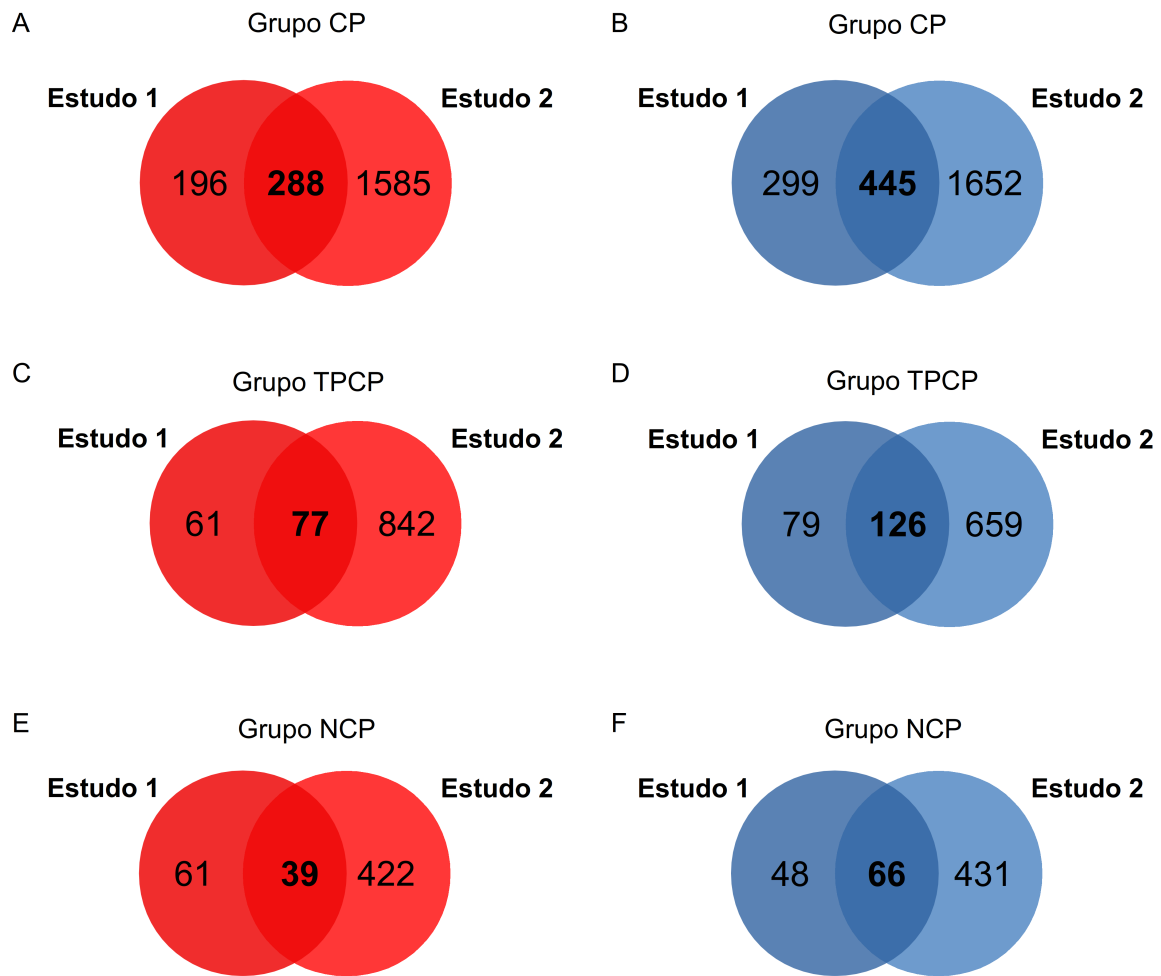


Figura 16. Transcritos diferencialmente expressos por categoria. (A) Número de transcritos CP superexpressos nos Estudos 1 e 2. (B) Número de transcritos CP subexpressos nos Estudos 1 e 2. (C) Número de transcritos TPCP superexpressos nos Estudos 1 e 2. (D) Número de transcritos TPCP subexpressos nos Estudos 1 e 2. (E) Número de transcritos NCP superexpressos nos Estudos 1 e 2. (F) Número de transcritos NCP subexpressos nos Estudos 1 e 2. Número de TDE em comum aos Estudos no centro do Diagrama de Venn e número de TDE únicos nas extremidades. CP, codificadores de proteínas. NCP, não codificadores de proteínas. TDE, transcritos diferencialmente expressos. TPCP, transcritos processados originados de genes codificadores de proteínas.

A análise em nível de transcritos permitiu observar uma discrepância entre o número de genes e de transcritos dentre os TDE em comum aos dois estudos e que são originados de genes codificadores de proteínas (Figura 17). Apenas 284 genes são responsáveis pela transcrição dos 365 transcritos superexpressos (288 CP + 77 TPCP) e 405 genes dão origem aos 571 transcritos subexpressos (445 CP + 126 TPCP). A maior parte dos genes teve apenas transcritos codificadores de proteínas diferencialmente expressos (226 genes superexpressos e 317 genes subexpressos), outros apresentaram transcritos de mais de uma categoria incluindo pelo menos um do tipo codificador de proteína (29 genes superexpressos e 51 genes

subexpressos). Interessante observar que alguns genes codificadores de proteínas deram origem apenas a transcritos do grupo TPCP (29 genes superexpressos e 37 subexpressos).

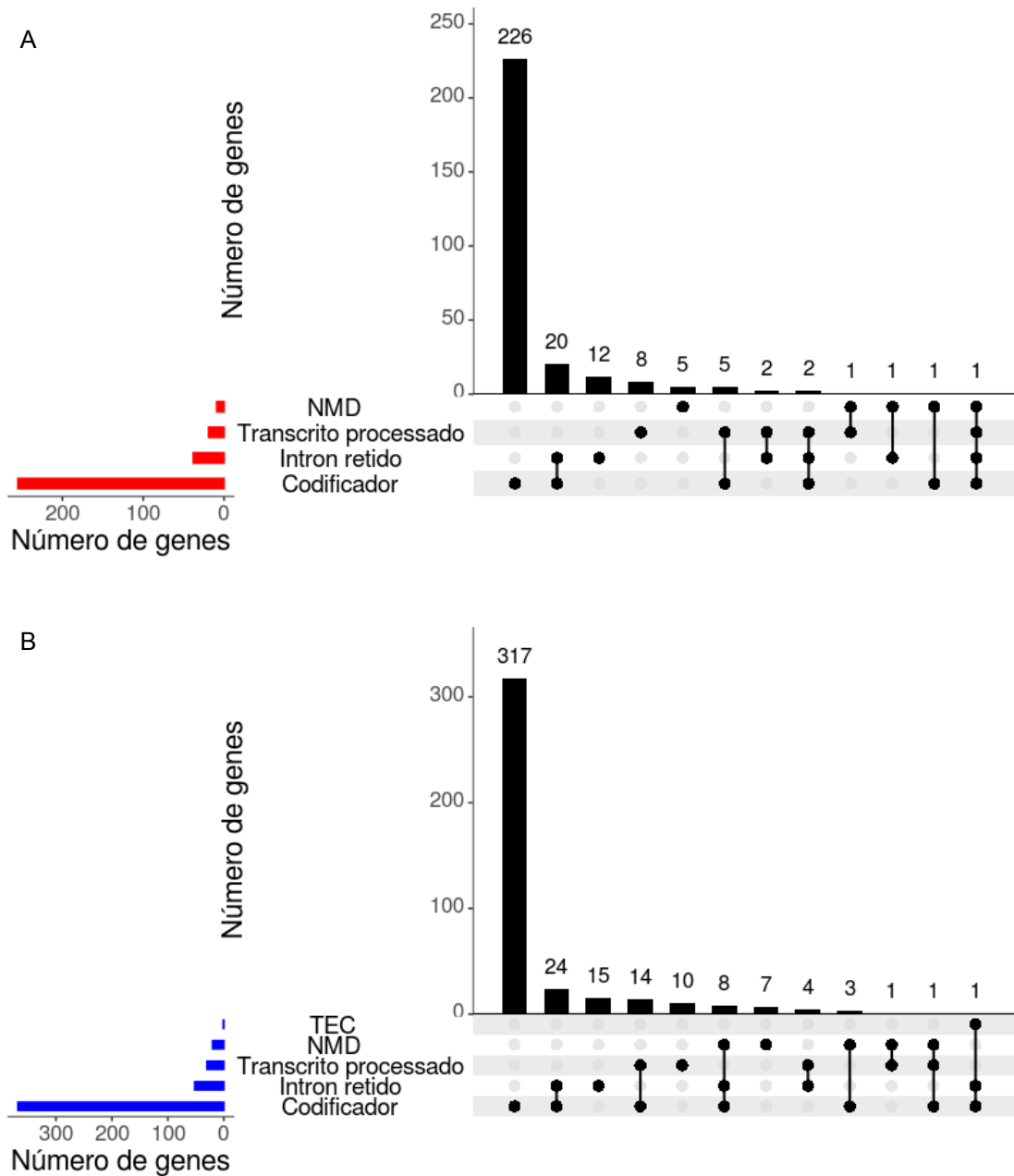


Figura 17. Distribuição de genes por categorias de transcritos. (A) Genes com transcritos superexpressos em comum ao Estudo 1 e Estudo 2. (B) Genes com transcritos subexpressos em comum ao Estudo 1 e Estudo 2. Valores acima das barras pretas representam o número de genes contidos nas categorias de transcritos assinaladas por ponto preto. NMD, *nonsense mediated decay*. TEC, *to be experimentally confirmed*.

Dentre os genes que apresentaram mais de uma isoforma de transcrito diferencialmente expressa estão os genes *Kim-1*, *Mdm2* e *Mate1*. Estes genes serão utilizados como exemplo das possíveis modificações estruturais em diferentes isoformas de um mesmo gene.

O gene *Kim-1* apresentou duas isoformas de transcritos superexpressas devido à lesão causada por cisplatina, ambas codificadoras de proteína (identificadores Ensembl: ENSMUST00000047568 e ENSMUST00000109223). A diferença entre as duas isoformas está tanto no sítio de início da transcrição quanto no uso de éxons. O transcrito ENSMUST00000047568 possui 8 éxons, todos codificadores. Este mRNA possui 1.892 pb de comprimento e codifica uma proteína de 305 aminoácidos. Já o transcrito ENSMUST00000109223, possui 8 éxons, sendo 7 deles codificadores, seu mRNA possui 1.982 pb de comprimento e codifica uma proteína de 282 aminoácidos. Em relação ao transcrito ENSMUST00000047568, um éxon não codificador é adicionado ao início do transcrito ENSMUST00000109223 e o éxon codificador 4 é perdido.

A inserção de um éxon no início do transcrito ENSMUST00000109223 não altera o sítio de início da tradução, que neste transcrito se localiza no éxon 2, enquanto no transcrito ENSMUST00000047568 se localiza no éxon 1. A diferença no número de aminoácidos codificados nesses mRNAs é devido à perda do éxon 4 no transcrito ENSMUST00000109223. A perda do éxon 4 não altera a fase de leitura do transcrito e também não afeta o domínio de imunoglobulina localizado no segundo éxon codificador do gene *Kim-1* nem a região transmembrana desta proteína codificada no éxon 6. No entanto, a perda do éxon 4 afeta a região extracelular da proteína KIM-1, fazendo com que a mesma perca 23 aminoácidos numa região não estruturada.

O gene *Mdm2* apresentou 5 diferentes isoformas de transcritos superexpressas. O gene *Mdm2* possui 12 éxons com a porção codificadora nos éxons 3 a 12. Enquanto o éxon 12 codifica o domínio RING, responsável pela atividade ubiquitina ligase, os éxons 4 e 5 codificam a região responsável pela interação com p53. Os tipos de transcritos superexpressos neste trabalho foram codificador de proteína (ENSMUST00000020408), NMD (ENSMUST00000155285), íntron retido (ENSMUST00000131627 e ENSMUST00000132277) e transcrito processado (ENSMUST00000147823).

Dentre as isoformas supracitadas existem diferenças relacionadas à presença e ausência de éxons, no entanto somente as isoformas codificadoras de proteína NMD possuem região codificadora. A isoforma codificadora de proteína possui a região codificadora do gene íntegra. Já a isoforma NMD, por possuir um códon de parada precoce, não codifica a região

do domínio RING. Além disso, os mRNAs do tipo NMD são degradados, conforme apresentado no tópico “1.8 Transcritos originados de genes codificadores de proteínas”. Dessa maneira, a despeito da superexpressão de 5 isoformas distintas, apenas uma das isoformas é codificadora de proteína e, portanto, apta à tradução da proteína MDM2.

O gene *Mate1* possui 4 isoformas de transcrito anotadas atualmente, sendo duas do tipo codificador de proteína, uma do tipo NMD e uma do tipo transcrito processado. No presente trabalho, apenas as isoformas NMD (ENSMUST00000131723) e transcrito processado (ENSMUST00000147583) foram subexpressas após o tratamento com cisplatina. Por serem do grupo TPCP, essas isoformas de transcrito de *Mate1* não possuem região codificadora e, com isso, não estão aptas a serem traduzidas.

Os genes do grupo NCP também foram expressos em diferentes isoformas de transcrito, sendo que 84 genes foram responsáveis pela origem dos 105 TDE. Desses, 34 genes originam os 39 transcritos superexpressos e 50 genes originam os 66 transcritos subexpressos.

4.3.2 Transcrição *readthrough*

Os eventos de transcrição *readthrough* assim como seus níveis de expressão foram acessados através do programa DoGFinder para as amostras do Estudo 2. O programa DoGFinder identifica as regiões à jusante de genes anotados que possuem *reads* alinhadas.

O número total de DoGs encontrado, considerando-se aqueles que constam em todas as amostras do grupo controle e ou do grupo tratado, foi 1525. Desses, 655 (42,95%) ocorreram em todos os animais, independentemente do tratamento. Foram encontrados 880 (57,70%) DoGs comuns a todas as amostras do grupo controle e 1300 (85,25%) DoGs comuns a todas as amostras do grupo tratado, sendo 225 (14,75%) DoGs exclusivos das amostras de animais não tratados e 645 (42,30%) DoGs exclusivos de animais tratados com cisplatina (Figura 18). Dessa forma, tem-se que o tratamento com cisplatina induz um aumento de 2,87 vezes no número de DoGs exclusivos em relação aos animais não tratados.

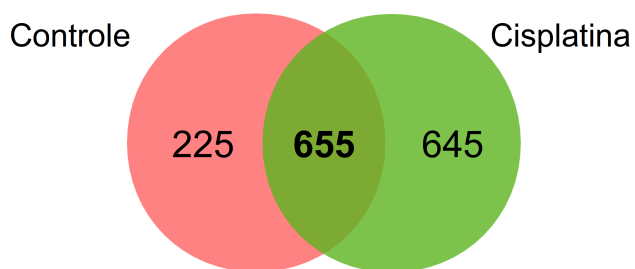


Figura 18. Eventos de transcrição *readthrough*. Número de DoGs que ocorrem exclusivamente em animais não tratados com cisplatina (rosa), que ocorrem exclusivamente em animais tratados com cisplatina (verde) e que ocorrem nas duas condições (interseção). Todas as amostras referem-se ao Estudo 2. DoGs, *Downstream Of Genes containing transcripts*.

Assim como o número de DoGs, o tamanho dos mesmos também aumenta devido ao tratamento. Considerando os 655 DoGs comuns a todas as amostras, o tamanho mediano dos DoGs dos animais não tratados é 9.800 pb, enquanto nos animais tratados a mediana do tamanho é 12.000 pb. Considerando os DoGs exclusivos de cada condição, temos que o tamanho mediano dos 225 DoGs exclusivos dos animais não tratados com cisplatina é de 8.200 pb. Enquanto o tamanho mediano dos 645 DoGs exclusivos dos animais tratados com cisplatina é de 9.082 pb.

O tratamento com cisplatina aumentou a expressão de aproximadamente 70% dos DoGs em relação aos animais não tratados, considerando os DoGs que constam em todas as amostras do grupo controle e ou do grupo tratado. Entretanto, para cerca de 30% dos DoGs, houve diminuição da expressão nos animais tratados (Figura 19). Sessenta e dois genes com transcritos superexpressos sofreram transcrição *readthrough*, sendo todos os DoGs gerados também superexpressos. Cento e vinte e dois genes com transcritos subexpressos sofreram transcrição *readthrough*, sendo que 112 DoGs foram subexpressos e 10 superexpressos.

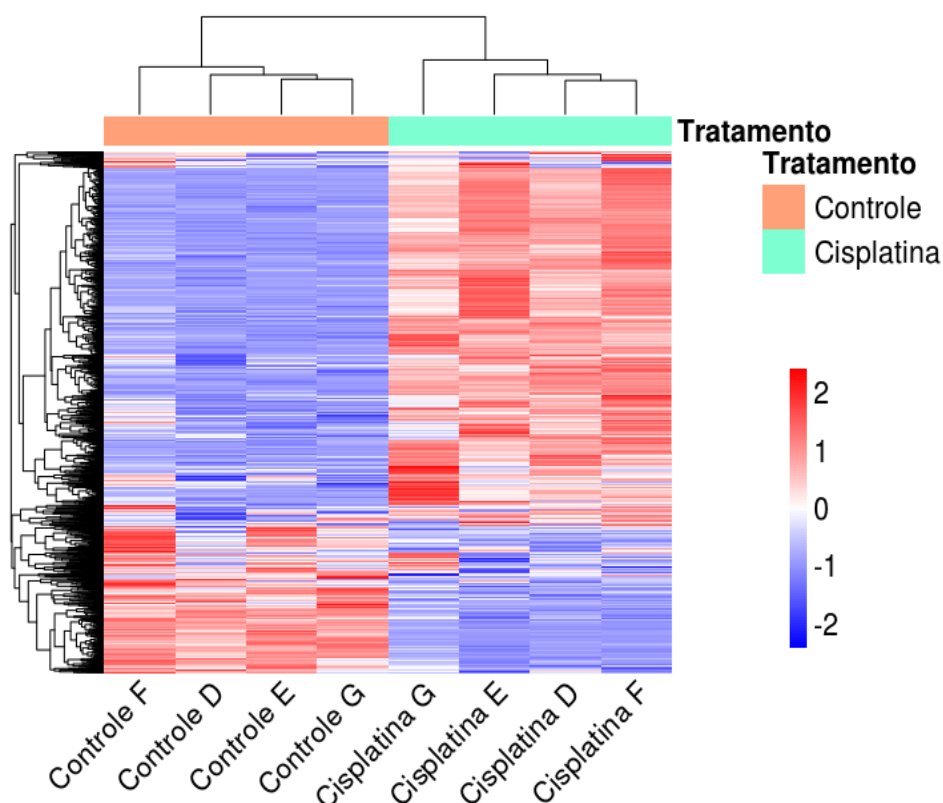


Figura 19. Heatmap dos níveis de expressão dos DoGs. DoGs presentes em todas as amostras do grupo controle e todas as amostras do grupo tratado. Todas as amostras referem-se ao Estudo 2. Escala de cor representa o *z-score* do RPKM dos DoGs. DoGs, *Downstream Of Genes containing transcripts*.

4.3.3 Análise funcional

Para os TDE dos grupos CP e TPCP e para os genes que sofreram transcrição *readthrough*, a análise funcional foi realizada através do enriquecimento de processos biológicos definidos pelo Gene Ontology usando o programa topGO. Foram considerados significativos aqueles termos GO com valor-p menor que 0,01. Já no caso dos TDE do grupo NCP, a análise funcional foi realizada através de pesquisa bibliográfica e das informações disponíveis no banco de dados NONCODE.

4.3.3.1 *Transcritos superexpressos*

Dentre os transcritos codificadores de proteínas (grupo CP), os processos biológicos enriquecidos convergem, especialmente, para a ocorrência de apoptose pelas vias intrínsecas,

enquanto a via extrínseca da apoptose é regulada negativamente. A proliferação e diferenciação celular aparecem como processos enriquecidos sendo o primeiro regulado negativamente e o segundo, positivamente, acompanhado de regulação positiva do desenvolvimento renal. A resposta ao dano do DNA aparece como um processo à parte e também em outros processos mais específicos como, por exemplo, resposta à luz ultravioleta e à radiação gama, que, assim como a cisplatina, formam adutos e causam quebra da dupla fita de DNA. Também foram enriquecidos processos biológicos relacionados à inflamação, à resposta ao estresse oxidativo, ao controle de ciclo celular e à transcrição de promotor de Pol II, dentre outros (Figura 20A e Apêndice D).

Os transcritos superexpressos do grupo TPCP apresentaram apenas processos biológicos relacionados à formação de microtúbulos com enriquecimento significativo. No limiar de significância, porém não significativos, estão os processos de regulação da sinalização da proteína quinase B e parada de ciclo celular, sendo o primeiro envolvido em diversos processos celulares como proliferação celular e apoptose (Figura 20B e Apêndice E).

À exceção do transcrito *Snhg15-204* (código Ensembl ENSMUST00000147762), não foram encontrados na pesquisa bibliográfica artigos referentes a transcritos específicos da categoria NCP. Foram encontrados apenas artigos relacionados ao gene de origem dos transcritos. Ainda assim, dos 34 genes não codificadores que dão origem aos 39 transcritos superexpressos, apenas para sete genes foram encontrados artigos relacionados, a saber, os lincRNAs *Gm26917*, *Gm26669*, *Gm17501*, *Mirt1*, *Trp53cor1*, também conhecido como *lincRNA-p21*, o transcrito processado *Snhg15* e o antissenso *1700110K17Rik*. Importante observar que a pesquisa bibliográfica não se limitou ao camundongo, espécie utilizada nos estudos abordados, nem ao tipo celular ou órgão, nem mesmo ao tratamento.

Assim como nos transcritos codificadores de proteínas, a apoptose foi uma função recorrente dentre os genes não codificadores superexpressos que tiveram suas funções estudadas. A apoptose é regulada positivamente pelos lincRNAs *Mirt1* (LI; ZHOU; HUANG, 2017; LIU et al., 2019a) e *lincRNA-p21* (HUARTE et al., 2010). *Mirt1* também foi relacionado à regulação positiva da inflamação (LI; ZHOU; HUANG, 2017; LIU et al., 2019a) e ao remodelamento cardíaco em infarto de miocárdio (ZANGRANDO et al., 2014). Já o lincRNA *Gm26917* foi relacionado à proteção contra a apoptose em células-tronco do músculo esquelético (CHEN et al., 2018b).

Outro processo biológico também encontrado nos genes não codificadores de proteínas é o estresse oxidativo. O lincRNA *Gm26669* pode estar envolvido no aumento do estresse oxidativo e na disfunção mitocondrial (XIAO et al., 2019). O lincRNA *Gm17501* e o

antisense *1700110K17Rik* foram estudados a partir de redes de coexpressão entre RNAs codificadores e não codificadores e compartilharam o módulo com genes codificadores sabidamente envolvidos no desenvolvimento de hipertrofia cardíaca (ZHANG et al., 2018b). O gene *Snhg15* é largamente estudado em diversos tipos de câncer e é descrito por estar envolvido na proliferação, migração, invasão, parada de ciclo celular e apoptose (TONG et al., 2019), sendo primeiramente caracterizado por estar relacionado ao estresse celular induzido por substâncias químicas (TANI; TORIMURA, 2013; TANI et al., 2017). O transcrito ENSMUST00000147762 do gene *Snhg15* pode estar envolvido na regulação da ativação de AMPK induzida por isquemia do miocárdio (SUN et al., 2019).

O banco de dados NONCODE possui predição de função, através da atribuição de termos GO, para 310 genes não codificadores de proteínas de camundongo. Entretanto, apenas um desses genes foi superexpresso devido ao tratamento com cisplatina, sendo ele o lincRNA *3300005D01Rik*. Dentre as funções preditas atribuídas a este gene estão a regulação negativa da apoptose e regulação da sinalização celular e da proteólise.

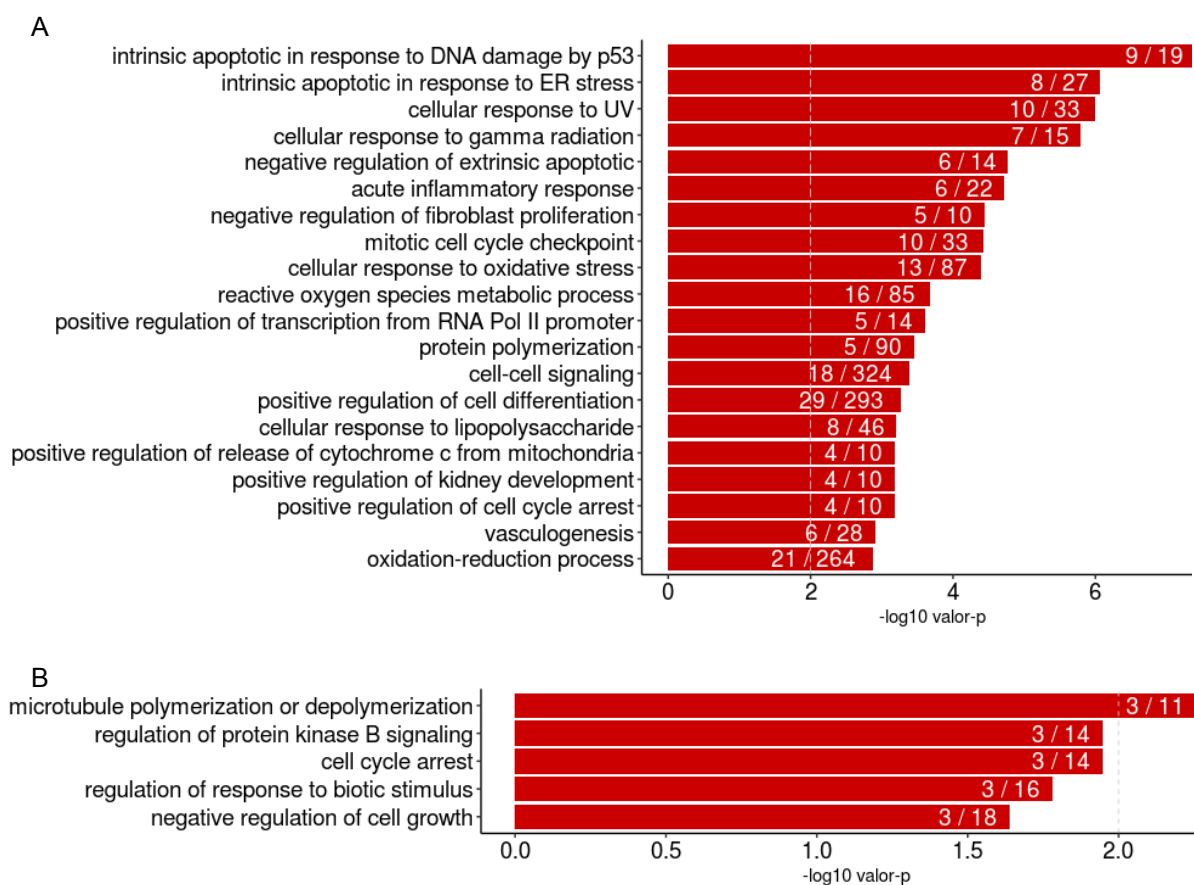


Figura 20. Processos biológicos enriquecidos dentre os genes codificadores de proteína superexpressos. (A) Processos biológicos enriquecidos dentre os transcritos codificadores de proteína (grupo CP). (B) Processos biológicos enriquecidos dentre os transcritos processados originados de genes codificadores de proteína (grupo TPCP). Linha pontilhada, valor de corte de significância estatística, $p = 0.01$. Valores dentro das barras indicam o número de genes significativos e o número de genes anotados dentro de cada processo biológico.

4.3.3.2 *Transcritos subexpressos*

Assim como para os transcritos superexpressos, o enriquecimento de processos biológicos foi realizado separadamente para os transcritos subexpressos do grupo CP e do grupo TPCP. Porém, de maneira oposta, há grande semelhança entre os processos biológicos enriquecidos em ambos os grupos (Figura 21 e Apêndices F e G).

Dentre os processos em comum, estão a biossíntese e o catabolismo de aminoácidos, processo de óxido-redução, catabolismo de drogas e transporte de íon sódio. Mesmo os processos biológicos exclusivos de cada grupo mantêm um certo grau de relacionamento seja com os processos em comum ou com os processos exclusivos. No grupo CP, dentre os processos exclusivos estão relacionados ao transporte de drogas e de ânions, catabolismo de cofator, biossíntese de ácidos graxos e homeostase de lipídeos. Já no grupo TPCP, aparecem termos relacionados ao catabolismo de lipídeos e à homeostase de íons.

A pesquisa bibliográfica pelo nome ou código dos transcritos subexpressos da categoria NCP não retornou nenhum resultado, havendo resultados apenas para as buscas com o nome dos genes. Dos 50 genes não codificadores de proteínas que dão origem aos 66 transcritos subexpressos, apenas 4 retornaram resultado positivo na pesquisa bibliográfica na base de dados PubMed, sendo eles os pseudogenes *Gpr137b-ps* e *Gm6300*, o lincRNA *Gm5524* e o antissenso *Airn*. O pseudogene *Gpr137b-ps* foi descrito por estar envolvido na regulação positiva da angiogênese, sendo que a supressão do mesmo leva à diminuição da proliferação celular e da neoangiogênese (WANG et al., 2019). Já o pseudogene *Gm6300*, subexpresso em rins de camundongos diabéticos, ainda não teve sua função esclarecida (KOMERS et al., 2014).

O lincRNA *Gm5524* foi estudado em camundongos com nefropatia diabética e sua função foi analisada em conjunto com a do pseudogene *Gm15645*. A superexpressão do primeiro associada à subexpressão do segundo mostrou-se envolvida na regulação de processos apoptóticos (FENG et al., 2018). Ressalta-se que o pseudogene *Gm15645* não figurou na lista de genes diferencialmente expressos deste trabalho, limitando a inferência de função do *Gm5524* no contexto aqui estudado. O gene *Airn*, antissenso do gene codificador de proteína *Igf2r*, atua como silenciador deste último e também dos genes *OCT2* e *OCT3* (*SLC22A3*) (KANDURI, 2016).

Não foram encontradas funções previstas para os transcritos subexpressos no banco de dados NONCODE.

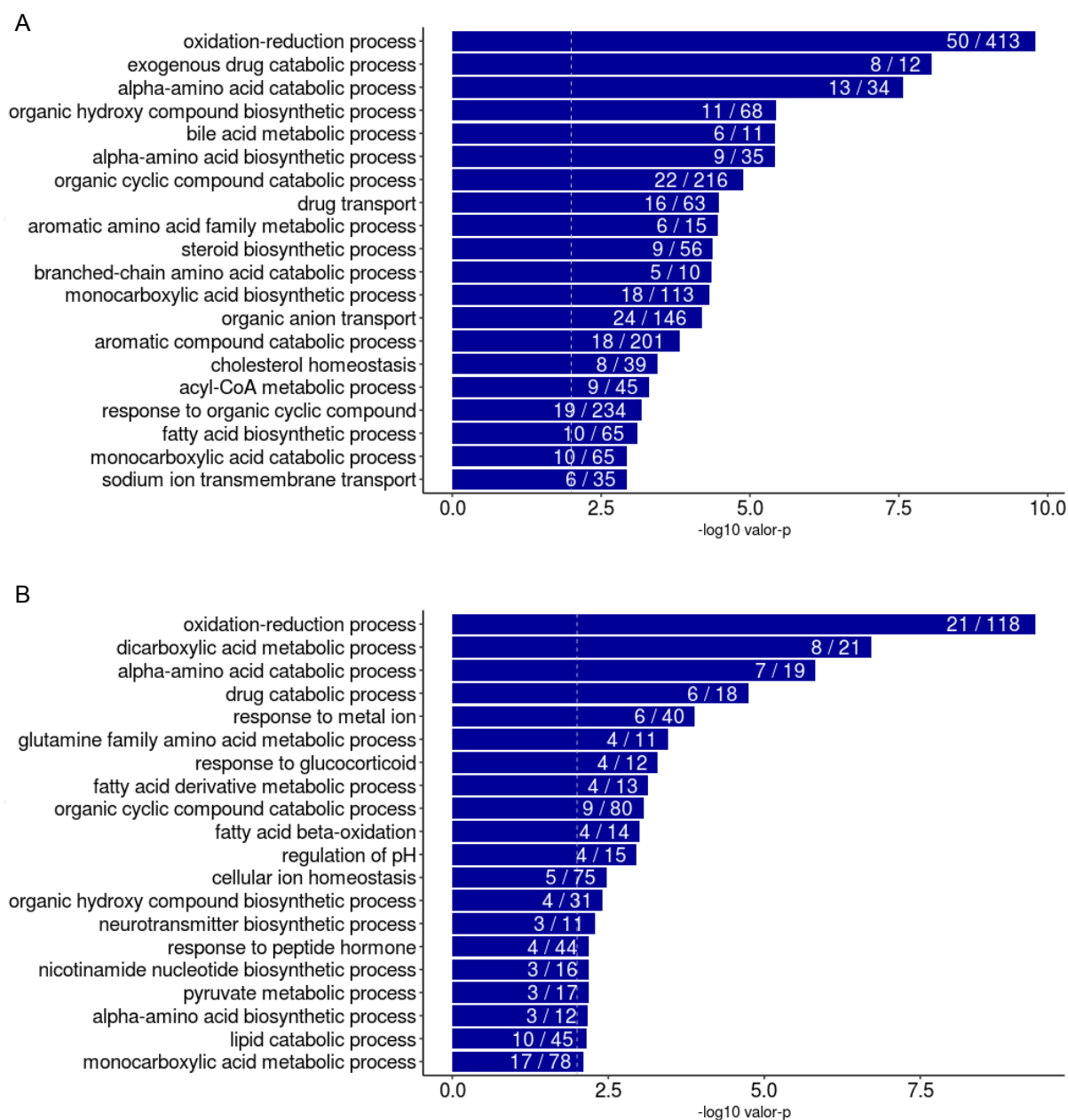


Figura 21. Processos biológicos enriquecidos dentre os genes codificadores de proteína subexpressos. (A) Processos biológicos enriquecidos dentre os transcritos codificadores de proteína (grupo CP). (B) Processos biológicos enriquecidos dentre os transcritos processados originados de genes codificadores de proteína (grupo TPCP). Linha pontilhada, valor de corte de significância estatística, $p = 0.01$. Valores dentro das barras indicam o número de genes significativos e o número de genes anotados dentro de cada processo biológico.

4.3.3.3 Genes que sofreram transcrição *readthrough*

A fim de identificar os processos biológicos enriquecidos dentre os genes que sofreram transcrição *readthrough*, a análise foi realizada separadamente para os genes cujos DoGs tiveram expressão aumentada ou diminuída (Figura 22, Apêndices H e I). Ressalta-se que a

análise funcional não foi realizada com os DoGs, mas com os genes em que a transcrição *readthrough* os gerou.

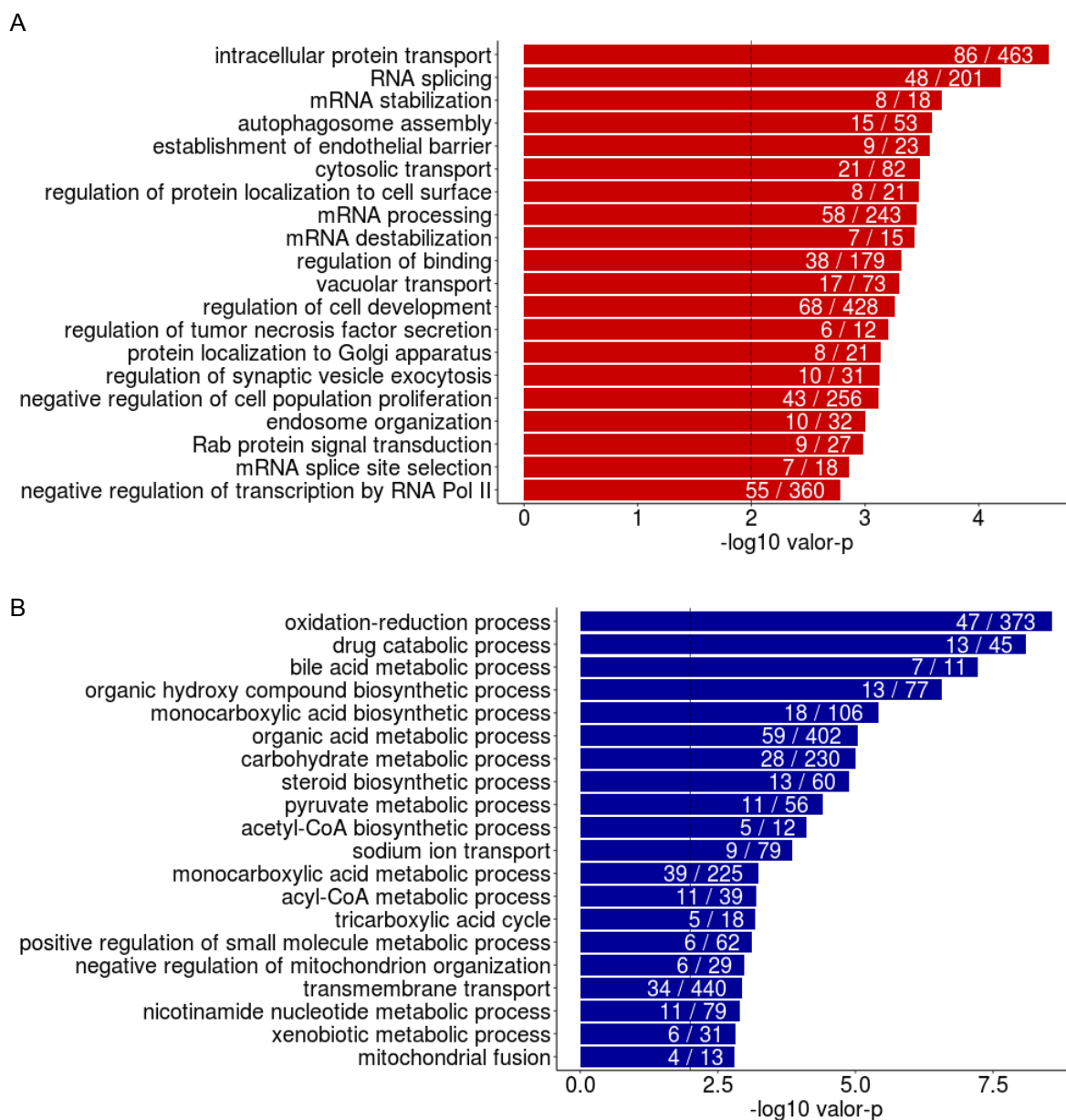


Figura 22. Processos biológicos enriquecidos dentre os genes que sofreram transcrição *readthrough*. (A) Processos biológicos enriquecidos dentre os genes que deram origem aos DoGs superexpressos. (B) Processos biológicos enriquecidos dentre os genes que deram origem aos DoGs subexpressos. Linha pontilhada refere-se ao ‘valor de corte de significância estatística, $p = 0.01$. Valores dentro das barras indicam o número de genes significativos e o número de genes anotados dentro de cada processo biológico. DoGs, *Downstream Of Genes containing transcripts*.

Dentre os genes em que os DoGs aumentaram de expressão, houve grande enriquecimento para o processamento de mRNA e transporte e localização de proteínas. Dentre os genes em que os DoGs diminuíram a expressão após o tratamento com cisplatina, os principais processos biológicos enriquecidos estão relacionados ao metabolismo basal.

5 DISCUSSÃO

No presente trabalho, foram analisadas amostras de sequenciamento de RNA de rins de camundongos tratados com cisplatina originados de dois estudos prévios disponibilizados no banco de dados SRA e que podem ser acessados através do banco de dados GEO pelos números de acesso GSE69652 e GSE106993. Em ambos, os autores tiveram como foco principal analisar pré-tratamento ou condicionamentos que mitigassem ou anulassem o efeito nefrotóxico da cisplatina.

No trabalho GSE69652, aqui denominado “Estudo 1”, Galgamuwa e colaboradores (GALGAMUWA et al., 2016) avaliaram se o pré-tratamento com dicloroacetato seria capaz de diminuir os efeitos nefrotóxicos da cisplatina. Os autores concluíram que a administração prévia de dicloroacetato confere proteção contra os danos renais causados pela cisplatina sem prejudicar seu efeito antitumoral. Essa proteção ocorre através da manutenção da respiração celular, aumento da proliferação de células tubulares e da inibição da morte celular.

No trabalho GSE106993, aqui denominado “Estudo 2”, Späth e colaboradores (SPÄTH et al., 2019) analisaram se os condicionamentos de hipóxia e restrição calórica previnem a nefrotoxicidade induzida por cisplatina a partir de uma abordagem proteômica associada à transcriptômica. Os dois condicionamentos foram capazes de reduzir os danos renais causados pela cisplatina, especialmente a restrição calórica, que em diversas análises apresentou padrões similares aos dos animais não tratados com cisplatina. Os autores concluíram que a cisplatina pode induzir a lise de membrana celular gerando um ambiente inflamatório e morte celular e a restrição calórica promoveu síntese de ácidos graxos, prevenindo a perda da integridade da membrana.

Adicionalmente às metodologias empregadas e resultados obtidos pelos autores dos dois estudos supracitados, optamos por realizar uma análise do transcriptoma mais profunda utilizando os mesmos dados. No presente trabalho, analisamos o dano causado pela cisplatina através de mudanças nos padrões de expressão gênica ao nível dos transcritos, com foco em isoformas geradas pelo mecanismo de *splicing* alternativo, buscando integrar informações de genes codificadores e não codificadores de proteínas nos mecanismos envolvidos na nefrotoxicidade induzida por cisplatina. Além de avaliar o efeito desse medicamento no término da transcrição através da expressão de DoGs.

Os Estudos 1 e 2 utilizaram-se de estratégias distintas para o tratamento dos animais e para o sequenciamento de RNA. A análise de expressão dos transcritos comuns aos dois

estudos implica na perda parcial de informações relacionadas aos transcritos que são diferencialmente expressos exclusivamente em um ou em outro estudo. Por outro lado, essa abordagem revela mecanismos de nefrotoxicidade consistentes, uma vez que a cisplatina foi capaz de induzi-los independentemente do tempo de exposição, da linhagem de camundongo utilizada ou quaisquer outras diferenças nos desenhos experimentais.

A baixa sensibilidade e especificidade dos marcadores atuais de dano renal fomentam a busca por novos biomarcadores. Nesse sentido, o FDA aprovou novas moléculas para uso diagnóstico e prognóstico em animais e, caso a caso, em humanos e, dentre elas, está a proteína KIM-1. A expressão de KIM-1 é induzida por danos renais sendo possível detectar aumento no mRNA, proteína tecidual e proteína urinária. Os níveis de KIM-1 se elevam rapidamente após tratamento com medicamentos nefrotóxicos em animais (VAIDYA et al., 2010; SPÄTH et al., 2019), humanos (LIAO et al., 2019), linhagens celulares animais (GROSSI et al., 2017; SILVA et al., 2017) e humanas (CAMPOS et al., 2018). KIM-1 também é induzida por outros danos renais como isquemia (ICHIMURA et al., 1998) e necrose tubular aguda (HAN et al., 2002). Além de se mostrar um biomarcador superior aos tradicionalmente utilizados na clínica, KIM-1 confere às células tubulares renais características fagocíticas que permitem a internalização de corpos apoptóticos e debris celulares em rins lesionados (ICHIMURA et al., 2008).

No presente trabalho, duas isoformas de transcritos de *Kim-1* foram encontradas superexpressas devido à lesão causada por cisplatina, ambas codificadoras de proteína (ENSMUST00000047568 e ENSMUST00000109223). A diferença entre essas isoformas de transcritos que afeta a proteína relaciona-se à perda do éxon 4 que culmina na perda de 23 aminoácidos na região extracelular de KIM-1, como descrito anteriormente. Não foi possível determinar se essa modificação altera de alguma maneira a função da proteína. É possível que nenhuma alteração significativa ocorra, uma vez que os aminoácidos perdidos localizam-se em uma região não estruturada. No entanto, é necessário que outros estudos sejam realizados comparando-se as estruturas e as funções das proteínas geradas por essas isoformas de transcritos.

Além de *Kim-1*, outro biomarcador foi superexpresso após tratamento com cisplatina, *Lcn2* (NGAL). NGAL apresenta expressão aumentada frente a diferentes danos renais, como lesão de isquemia e reperfusão (SUPAVEKIN et al., 2003), medicamentos nefrotóxicos (WANG et al., 2008a; SPÄTH et al., 2019) e LRA pós-cirúrgica (YONEYAMA et al., 2019). Este biomarcador também se mostrou mais consistente que IL-18 e cistatina C para

diagnóstico de LRA em pacientes com câncer em tratamento com cisplatina (BUNEL et al., 2017).

O número de estudos focados na nefrotoxicidade induzida por medicamentos utilizando dados oriundos de sequenciamento de RNA ainda é escasso. Mesmo a lesão renal aguda, condição clínica induzida por diversos fatores, inclusive medicamentos, ainda é pouco estudada utilizando essa metodologia. Independentemente da metodologia utilizada para acessar a expressão gênica, a identificação de miRNAs que possuem potencial para uso como biomarcadores tem ganhado atenção, seja para fins de diagnóstico, prognóstico ou acompanhamento da evolução do dano renal. Já os lncRNAs começaram a ganhar espaço no estudo da LRA mais recentemente e ainda existem poucos estudos neste campo, sendo mais comumente encontrada a relação entre a expressão dos lncRNAs e o dano renal do que a função, mecanismos e alvos desses RNAs (BRANDENBURGER et al., 2018; REN et al., 2019). Poucos transcritos diferencialmente expressos do grupo NCP em rins de camundongos tratados com cisplatina em relação a animais não tratados possuem estudos que elucidem seu papel biológico. No entanto, mesmo que os estudos sobre os lncRNAs disponíveis não estejam diretamente relacionados à nefrotoxicidade, é possível inferir o papel de alguns deles dentro deste contexto.

Os transportadores de membrana têm papel essencial na nefrotoxicidade induzida por cisplatina por atuarem como porta de entrada desse medicamento nas células. O transportador OCT2 está diretamente envolvido na entrada da cisplatina pela face basolateral das células do túbulo proximal (FILIPSKI et al., 2009; CIARIMBOLI et al., 2010). *Airn* é um gene não codificador de proteína, antisseno ao gene *Igf2r*, envolvido no silenciamento dos genes codificadores de proteína *Igf2r*, *Oct2* e *Oct3*. *Airn* silencia o gene *Igf2r* de maneira distinta da qual silencia os genes *Oct*. Por ser sobreposto ao gene *Igf2r*, o próprio processo de transcrição de *Airn* impede a transcrição de *Igf2r*. Já o silenciamento dos genes *Oct2* e *Oct3* é feito pelo RNA de *Airn* através do recrutamento de modificadores de cromatina (LATOS et al., 2012; KANDURI, 2016; ANDERGASSEN et al., 2019).

Camundongos tratados com cisplatina apresentaram subexpressão renal de *Airn* em relação aos animais não tratados, enquanto *Oct2* não apresentou alteração significativa na comparação entre os dois grupos. Até o momento, a relação entre os danos renais causados pela cisplatina e a regulação de expressão de *Airn* sobre *Oct2* ainda não foi estudada. Sabendo que transportadores de cátion orgânico – OCTs – possuem expressão órgão-específica e que cisplatina interage com OCT2, mas não com outros membros dessa família (CIARIMBOLI, 2014), a indução de *Airn* pode ser um alvo terapêutico para o controle do principal

transportador renal de cisplatina. Potencialmente, a atuação sobre este alvo não afetará a entrada da cisplatina em tumores, sendo necessários estudos específicos para validação destas hipóteses.

Além dos transportadores que permitem a entrada da cisplatina nas células renais, aqueles que fazem o efluxo da molécula para fora das células também participam dos processos tóxicos. Camundongos *knockout* para o transportador de efluxo *Mate1*, têm maior concentração renal de cisplatina que camundongos que expressam normalmente este gene, favorecendo a potencialização da nefrotoxicidade (NAKAMURA et al., 2010). Apenas isoformas não codificadoras de proteínas foram subexpressas nos camundongos tratados com cisplatina. Não foi possível determinar como a subexpressão de isoformas não codificadoras afetaria a expressão proteica de MATE1. É possível que, ao diminuir a expressão de isoformas de transcritos não codificadores, os níveis de isoformas codificadoras são virtualmente aumentados, numa tentativa de aumentar o efluxo de cisplatina e diminuir a nefrotoxicidade.

Galgamuwa e colaboradores (GALGAMUWA et al., 2016) observaram que o pré-tratamento com dicloroacetato atenuou a sinalização de apoptose mediada por p53 induzida por cisplatina em rins de camundongos. Esse resultado foi acompanhado pela observação morfológica de redução da apoptose renal. No presente trabalho, o processo biológico mais significativo enriquecido dentre os transcritos superexpressos CP foi a apoptose em resposta ao dano de DNA mediada por p53.

p53 é um gene supressor de tumor que dispara processos celulares que culminam na parada do ciclo celular e apoptose. Esses processos são executados a partir da ativação e repressão de inúmeros genes por p53 (PABLA; DONG, 2008). Um dos alvos de p53 é o *lincRNA-p21* (*Trp53cor1*) que foi descrito e caracterizado pela primeira vez por Huarte e colaboradores (HUARTE et al., 2010) e recebeu este nome pela proximidade genômica com o gene *Cdkn1a*, conhecido por *p21*. Os autores identificaram que o promotor do *lincRNA-p21* é alvo de p53 e tem sua transcrição ativada por p53 após danos no DNA. *LincRNA-p21* recruta a ribonucleoproteína nuclear hnRNP-K que se liga aos promotores de genes a serem reprimidos como parte da resposta canônica de p53.

Wu e colaboradores (WU et al., 2014) identificaram que *lincRNA-p21* regula a atividade de p53 através da ligação com a proteína MDM2, sendo esse um outro mecanismo de ação de *lincRNA-p21*. A região central do *lincRNA-p21* se liga ao domínio RING localizado na porção C-terminal de MDM2, inviabilizando a atuação de MDM2 sobre p53. MDM2 é uma ubiquitina ligase que tem a transcrição ativada por p53. MDM2 regula os níveis de p53 através da ubiquitinação, o que resulta na degradação de p53 pelo proteassomo

(KUBBUTAT; JONES; VOUSDEN, 1997). Em suma, *lincRNA-p21* é transcricionalmente ativado por p53 e tem ação nuclear e citoplasmática, atuando tanto na regulação negativa de genes à jusante da via de sinalização de p53 como na manutenção dos níveis de p53 através da interação com MDM2 (Figura 23A).

O papel de *lincRNA-p21* tem sido estudado em condições como aterosclerose (HE et al., 2015), doença de Parkinson (XU et al., 2018), fibrose hepática (ZHENG et al., 2015; YU et al., 2017) e nefropatia diabética (ZHANG et al., 2018a). Entretanto, não causa surpresa o fato deste lncRNA, que participa da regulação do ciclo celular e da apoptose, ser mais amplamente estudado em câncer (WANG et al., 2014b, 2017; DING et al., 2017; ZHANG et al., 2019; ZHOU et al., 2019). A expressão de *lincRNA-p21* foi também identificada após estímulo tóxico. Em um estudo conduzido em camundongos tratados com furano, um composto carcinogênico, os autores sugeriram que os genes *lincRNA-p21* e *p21* podem ser transcritos bidirecionalmente no fígado devido ao estímulo desse xenobiótico (RECIO et al., 2013).

No presente trabalho, observou-se que os tanto *lincRNA-p21* quanto *p21* são superexpressos na lesão causada por cisplatina, entretanto não é possível afirmar se houve ou não transcrição bidirecional. O papel de *lincRNA-p21* na nefrotoxicidade ainda não foi elucidado. Tendo em vista que todos os genes aqui citados, reguladores ou regulados por *lincRNA-p21*, foram superexpressos nos camundongos tratados com cisplatina, pode-se inferir que *lincRNA-p21* participa ativamente da apoptose induzida por cisplatina.

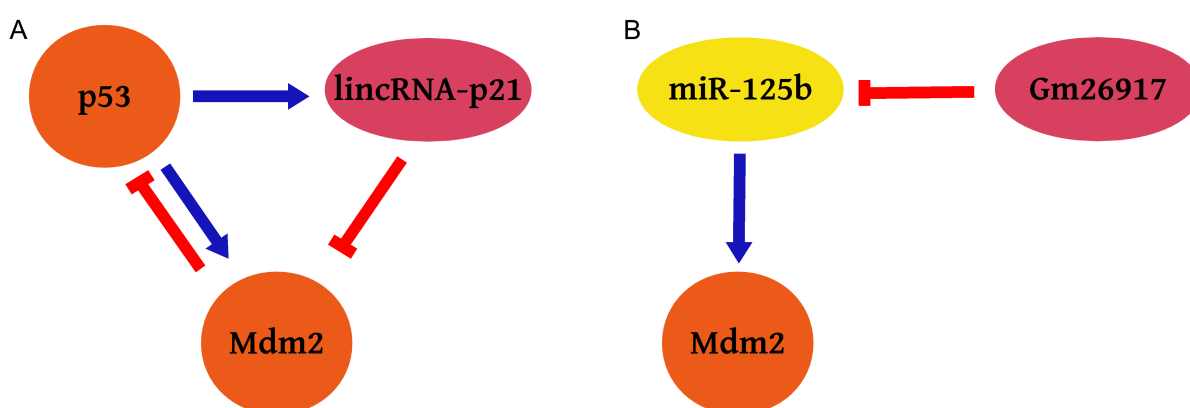


Figura 23. Mecanismos de regulação de Mdm2. (A) P53 induz a transcrição do lncRNA *lincRNA-p21* e de Mdm2. MDM2 regula negativamente os níveis de P53, enquanto *lincRNA-p21* liga-se a MDM2 impedindo sua atuação sobre P53. (B) O microRNA miR-125b induz Mdm2, entretanto o lncRNA Gm26917 pode atuar como esponja de miR-125b, impedindo sua ação sobre Mdm2. Círculo laranja, gene codificador de proteína. Elipse amarela, microRNA. Elipse rosa, lncRNA. lncRNA, RNA longo não codificador.

Diferentemente do *lincRNA-p21*, o *lincRNA Gm26917* não foi amplamente estudado até o momento, apenas um estudo investigou o papel deste gene. Chen e colaboradores (CHEN et al., 2018b) identificaram que *Gm26917* é necessário à sobrevivência de células satélites musculares através da interação com o miRNA pró-apoptótico miR-29b. Esses autores sugerem que *Gm26917* pode atuar como esponja de miRNA e que miR-125b é potencialmente um de seus alvos de interação. Altos níveis de miR-125b foram encontrados em ratos após lesão renal de isquemia e reperfusão, sendo um potencial biomarcador diagnóstico deste tipo de dano (GÜÇLÜ et al., 2017). Em 2013, Joo e colaboradores (JOO et al., 2013) mostraram que o mecanismo de ação de miR-125b leva à indução de *Mdm2* e protege contra os danos renais induzidos por cisplatina, uma vez que MDM2 atua como inibidor de p53. Sabendo que *Gm26917* é superexpresso em rins de camundongos tratados com cisplatina e considerando a predição de sua atuação como esponja do miR-125b, *Gm26917* contribui para a manutenção da apoptose através do bloqueio da ação de miR-125b em *Mdm2* (Figura 23B).

Em resposta ao dano de DNA induzido por radiação UV e cisplatina, o transcrito de *MDM2* sofre *splicing* alternativo de maneira que os éxons 4 a 11 são removidos e o éxon 3 é ligado diretamente ao éxon 12. Com isso, a proteína MDM2 derivada do *splicing* alternativo perde a capacidade de ligar-se a p53, apesar de manter o domínio RING (SINGH et al., 2009), sendo esse *splicing* alternativo de *MDM2* induzido pelo dano de DNA conservado entre camundongos e humanos (CHANDLER et al., 2006). Singh e colaboradores (SINGH et al., 2009) identificaram que a sequência que regula o *splicing* de *MDM2* de forma a manter todos os éxons está localizada no íntron 11. Uma sequência regulatória que também é conservada entre humanos e camundongos. No presente trabalho, foram identificadas 5 diferentes isoformas superexpressas de *Mdm2*, sendo que nenhuma delas corresponde à isoforma de *splicing* descrita acima. Das cinco isoformas de transcrito superexpressas de *Mdm2*, apenas uma é codificadora de proteína. Dessa forma, pode-se concluir que o número de mRNAs aptos à tradução não é equivalente ao nível transcricional do gene *Mdm2*.

Adicionando-se os mecanismos propostos para *lincRNA-p21* e *Gm26917* (Figura 23), tem-se que o aumento da expressão de *Mdm2* não leva, necessariamente, ao aumento dos níveis de proteínas funcionais, ou seja, o aumento de expressão de *Mdm2* induzido por cisplatina pode não levar a um aumento proporcional na degradação de p53. Em concordância com Chandler e colaboradores (CHANDLER et al., 2006), *p53* não sofreu *splicing* alternativo em decorrência da exposição à cisplatina.

Sabe-se que a apoptose induzida pela via de receptores de morte é um processo que ocorre em células tubulares em resposta à cisplatina (SERVAIS et al., 2008). Entretanto, o processo biológico de regulação negativa de sinalização de apoptose extrínseca via receptores de morte foi enriquecido dentre os transcritos superexpressos codificadores de proteína. Ao investigar quais genes contribuíram para o enriquecimento deste processo, observou-se que, na verdade, eram genes relacionados também a outros processos envolvidos na nefrotoxicidade induzida por cisplatina. Esses genes são *Hmox1*, *Tmbim1*, *Bcl2l1*, *Brcal*, *Fga* e *Icam1*.

Hmox1 contribuiu para o enriquecimento dos processos biológicos de resposta ao estresse oxidativo e resposta à substância tóxica. *Bcl2l1*, gene antiapoptótico, contribuiu para o enriquecimento do processo biológico de sinalização da apoptose intrínseca em resposta ao estresse do retículo endoplasmático e no *checkpoint* do ciclo celular. *Brcal* participa de processos metabólicos de espécies reativas de oxigênio e na transdução de sinais envolvida no dano de DNA. *Fga* e *Icam1* são genes conhecidos por participarem da resposta imune e também contribuíram para o enriquecimento do processo de polimerização de proteína. Apenas o gene *Tmbim1* contribuiu unicamente para o enriquecimento do processo biológico de regulação negativa da apoptose extrínseca.

Assim como muitos transcritos do grupo NCP, não se conhece o papel do pseudogene *Gpr137b-ps* na nefrotoxicidade. Este pseudogene é superexpresso e regula positivamente a angiogênese no cérebro de camundongos após estimulação de hipóxia. A supressão de *Gpr137b-ps* levou à diminuição da expressão de VEGF e de fosforilação de STAT3, sugerindo que a regulação positiva de ambos faz parte do mecanismo pelo qual *Gpr137b-ps* induz a angiogênese, processo este que envolve a indução de proliferação celular (WANG et al., 2019). Após o tratamento com cisplatina, o gene *Gpr137b-ps* foi subexpresso e o processo biológico de regulação negativa da proliferação celular foi enriquecido entre TDE superexpressos codificadores de proteína. Ainda que não se possa relacionar diretamente o papel regulatório de *Gpr137b-ps* sobre os genes VEGF e STAT3, sua subexpressão é condizente com a regulação negativa da proliferação celular encontrada nos rins dos camundongos tratados com cisplatina. Além disso, isoformas de transcritos codificadores e íntron retido de *Vegfa* e *Stat3* ficaram próximas ao valor de corte para serem consideradas subexpressas.

O lincRNA *Mirt1* foi identificado em coração de camundongos que sofreram infarto do miocárdio induzido e foi relacionado ao remodelamento cardíaco (ZANGRANDO et al., 2014). Estudos subsequentes identificaram que *Mirt1* é superexpresso após infarto agudo do

miocárdio e que o *knockdown* de *Mirt1* em camundongos nesta condição e em ratos com lesão cardíaca de isquemia e reperfusão levou à diminuição da ativação da via de sinalização NF- κ B, levando à diminuição de apoptose, inflamação e estresse oxidativo (LI; ZHOU; HUANG, 2017; LIU et al., 2019a). Outro gene não codificador possivelmente relacionado à via de sinalização NF- κ B é o lincRNA *Gm26669*.

Ao estudar o papel de genes não codificadores de proteína no efeito antifibrótico renal de norcantaridina, um antitumoral utilizado em câncer do trato digestório, Xiao e colaboradores (XIAO et al., 2019) observaram que *Gm26669* está relacionado à resposta inflamatória via NF- κ B e estresse oxidativo. As consequências da ativação de NF- κ B dependem do contexto de sua ativação, podendo regular a expressão de genes envolvidos na proliferação celular, apoptose, inflamação e resposta ao estresse (PERKINS, 2007). Na nefrotoxicidade induzida por cisplatina, a via de NF- κ B tem sido relacionada à resposta inflamatória (MILLER et al., 2010; MANOHAR; LEUNG, 2018). Até o momento, *Mirt1* tem sido estudado exclusivamente em lesões cardíacas e *Gm26669* teve seu papel analisado apenas no efeito antifibrótico de norcantaridina. Neste trabalho, mostramos que *Mirt1* e *Gm26669* foram superexpressos em rins de camundongos tratados com cisplatina. Tendo em vista que ambos estão relacionados à via de NF- κ B ressalta-se a importância de estudos direcionados para melhor compreender o papel destes genes não codificadores no processo inflamatório renal induzido por cisplatina.

O tratamento com cisplatina induziu o aumento de expressão dos lncRNA *Gm17501* e *1700110K17Rik*. Esses lncRNA também foram superexpressos em corações de camundongos com hipertrofia cardíaca. Nessa condição, ambos compartilharam o módulo de coexpressão com genes codificadores de proteína envolvidos no desenvolvimento da hipertrofia cardíaca, dentre eles o gene *Cyp2j13*. Dentro deste módulo, tanto os lncRNA quanto os genes codificadores de proteína apresentaram aumento de expressão nos camundongos com hipertrofia cardíaca em relação àqueles que não apresentavam alterações cardíacas (ZHANG et al., 2018b).

Ao contrário do que foi demonstrado em hipertrofia cardíaca, o gene codificador *Cyp2j13* foi subexpresso em rins após tratamento com cisplatina. *Cyp2j13* é membro da subfamília do citocromo P450 CYP2J. Enzimas dessa família participam do metabolismo dos ácidos graxos ácido araquidônico (AA) e ácido linoleico além do metabolismo de xenobióticos, sendo expressas em diversos tecidos extra-hepáticos (MA et al., 1999; GRAVES et al., 2013). No presente trabalho, houve enriquecimento de processos biológicos relacionados à biossíntese de ácidos graxos e catabolismo de drogas dentre os TDE

subexpressos do grupo CP. Dos genes pertencentes à subfamília CYP2J, cinco foram subexpressos após a exposição à cisplatina, a saber *Cyp2j5*, *Cyp2j7*, *Cyp2j9*, *Cyp2j11* e *Cyp2j13*. Análises de expressão tecidual em camundongo revelaram que *Cyp2j5* é expresso em rins e fígado, *Cyp2j9* apresenta maior expressão em cérebro, pulmão e fígado, mas também apresenta alta expressão em rins (GRAVES et al., 2013), já *Cyp2j11* é altamente expresso em rins e coração enquanto *Cyp2j13* apresenta alta expressão renal em camundongos machos e baixa expressão em outros tecidos independentemente do sexo do animal (GRAVES et al., 2013, 2015).

Embora esta família esteja envolvida no metabolismo de ácidos graxos, os genes *Cyp2j5*, *Cyp2j9*, *Cyp2j11* e *Cyp2j13* não contribuíram para o enriquecimento do processo biológico de biossíntese desses ácidos e, conforme esperado, esses genes contribuíram para o enriquecimento dos processos biológicos relacionados ao catabolismo de drogas. Importante ressaltar que genes da família CYP2J, em suas isoformas codificadoras de proteína, tiveram sua expressão diminuída após a exposição à cisplatina, o que, à primeira vista, poderia induzir à interpretação equivocada de que a diminuição da expressão de enzimas metabolizadoras de drogas levou a uma não inativação da cisplatina permitindo, portanto, a indução de toxicidade renal. No entanto, o conhecimento atual revela que enzimas do citocromo P450 não metabolizam cisplatina e seu papel na nefrotoxicidade ainda precisa ser melhor compreendido (QUINTANILHA et al., 2017). Esse caso mostra a importância da análise contextual dos resultados de enriquecimento de processos biológicos, pois, apesar das enzimas do citocromo P450 participarem do metabolismo de drogas, não é a partir deste mecanismo que atuam na nefrotoxicidade induzida por cisplatina.

O AA é um ácido graxo poli-insaturado presente nas membranas citoplasmáticas de mamíferos. A metabolização do AA por enzimas do citocromo P450 gera ácidos eicosanoides que atuam de forma parácrina e autócrina em processos celulares como inflamação, hipertensão e transporte renal de eletrólitos (KROETZ; ZELDIN, 2002). As enzimas da família do citocromo P450 2J metabolizam o AA, atuando como epoxigenases na biossíntese de ácidos epoxieicosatrienóicos (EET). CYP2J5 e CYP2J11 são expressos nos túbulos contorcidos proximais e CYP2J5 também aparece nos dutos coletores, regiões onde os EETs afetam o transporte de fluidos e eletrólitos (MA et al., 1999; GRAVES et al., 2013).

Camundongos fêmeas *knockout* para *Cyp2j5* apresentaram aumento de pressão arterial em relação àquelas homozigotas positivas para o mesmo gene. Esse aumento de pressão arterial foi acompanhado por um aumento na taxa de transporte nos túbulos proximais e diminuição dos níveis de estrogênio circulante nestes animais. Entretanto, a ausência do gene

Cyp2j5 não alterou a biossíntese de EET, aparentemente por ter havido um aumento compensatório na expressão do gene *Cyp2j9*. Camundongos machos *knockout* para *Cyp2j5* não apresentaram alteração na pressão arterial nem na taxa de transporte tubular em relação aos camundongos sem alteração na expressão de *Cyp2j5* (ATHIRAKUL et al., 2008). O uso de análogo de EET reduziu a pressão arterial, diminuindo a reabsorção de sódio pelos rins em camundongos obesos (SCHRAGENHEIM et al., 2018) e em ratos machos espontaneamente hipertensos (KHAN et al., 2014). Em suma, a diminuição dos níveis de EET levou a um aumento da reabsorção de sódio nos rins, que foi diminuída pelo uso de um análogo de EET.

Por outro lado, a exposição à cisplatina levou a uma diminuição tanto da expressão de genes responsáveis pela biossíntese de EET quanto de genes responsáveis pelo transporte de íon sódio. A diminuição de reabsorção de sódio em ratos hipertensos se deu pela redução da expressão do canal de sódio epitelial (ENaC, do inglês *epithelial sodium channel*) (KHAN et al., 2014). O uso de cisplatina não alterou a expressão dos genes que compõem o ENaC, portanto não foi possível identificar uma relação de EET e ENaC na nefrotoxicidade induzida por cisplatina como a que ocorre na hipertensão. Porém, não se pode descartar uma possível relação entre a diminuição dos níveis de EET e o transporte de eletrólitos nos rins como parte do processo nefrotóxico da cisplatina, sendo necessários mais estudos para compreender este ponto.

O principal produto da metabolização do AA por CYP2J5, CYP2J9, CYP2J11 e CYP2J13 é o 14,15-EET (MA et al., 1999; GRAVES et al., 2013). Utilizando a linhagem de células de túbulo proximal de porcos LLC-PKc14, Chen e colaboradores (CHEN; CAPDEVILA; HARRIS, 2001) observaram que 14,15-EET inibiu a apoptose induzida pela retirada do soro fetal bovino do meio de cultura, por peróxido de hidrogênio (H₂O₂), pelo antitumoral etoposido e pelo excesso de ácido araquidônico livre. A metabolização de AA em 14,15-EET pelas enzimas do citocromo P450 impedem a sinalização para apoptose gerada pelo excesso de AA livre, além de inibir a apoptose pela ativação da via de sinalização PI3K/Akt.

Considerando que genes que codificam metabolizadores de AA a EET tiveram suas expressões reprimidas após exposição à cisplatina e que EETs são capazes de inibir a apoptose, que é um dos principais fenótipos renais induzidos por cisplatina, a retomada da função desses ácidos graxos é um possível alvo terapêutico para a atenuação da nefrotoxicidade. De fato, Khan e colaboradores (KHAN et al., 2013) demonstraram que o uso de análogo de EET oral atenuou a nefrotoxicidade induzida por cisplatina em ratos através da redução da apoptose, além de redução do estresse oxidativo, da inflamação e do estresse do

retículo endoplasmático. Esses autores também concluíram, a partir de experimentos *in vitro* com linhagens de células tumorais, que a nefroproteção do análogo de EET não comprometeu a ação antitumoral da cisplatina. O papel das enzimas da família CYP2J na nefrotoxicidade induzida por cisplatina não tem sido estudado e o papel dos EETs ainda foi pouco investigado. Diante do exposto, pode-se concluir que tanto a família CYP2J quanto os EETs contribuem efetivamente para toxicidade renal da cisplatina, sendo um potencial alvo terapêutico para a atenuação dos danos renais seja através do seu papel no controle do transporte renal de íons, seja na redução da morte de células renais.

A identificação da razão pela qual a cisplatina causa ampla subexpressão gênica é complexa, pois pode envolver diversos fatores. O comprometimento da maquinaria de transcrição na superexpressão de genes e mecanismos específicos ativados por cisplatina são condições que podem atuar em conjunto causando subexpressão gênica. Observou-se que genes envolvidos no metabolismo basal celular foram amplamente subexpressos, especialmente aqueles envolvidos no metabolismo de aminoácidos e de lipídeos.

Späth e colaboradores (SPÄTH et al., 2019) observaram que a cisplatina induz rompimento de membrana plasmática, o que contribui para a morte celular. Ao expor os camundongos ao condicionamento de restrição calórica, a síntese de ácidos graxos preveniu a perda da integridade da membrana plasmática. No presente trabalho, observou-se que transcritos subexpressos contribuíram para o enriquecimento dos processos de biossíntese de ácidos graxos e esteroides, bem como para homeostase de colesterol. Em conjunto, conclui-se que a indução de síntese de lipídeos de membrana plasmática constitui uma possibilidade de intervenção terapêutica.

Os processos biológicos enriquecidos pelos transcritos superexpressos do grupo CP correspondem às vias de lesão já conhecidas para cisplatina, sendo as principais a apoptose, inflamação, estresse oxidativo e do retículo endoplasmático e lesões de mitocôndria. Além disso, também correspondem aos principais achados dos artigos de origem dos dados utilizados, no que tange apenas aos camundongos tratados com cisplatina em relação aos não tratados (GALGAMUWA et al., 2016; SPÄTH et al., 2019).

Entretanto, ao agregar informações sobre os genes não codificadores, foi possível revelar possíveis mecanismos até então não descritos no campo da nefrotoxicidade induzida por cisplatina para processos biológicos anteriormente relacionados a essa condição. Ao falar de enriquecimento de processos biológicos, reitera-se aqui a importância da análise contextual. Uma vez que os genes podem estar envolvidos em diferentes processos, é de suma importância que os processos biológicos sejam avaliados dentro da condição estudada para

que possíveis equívocos sejam evitados. Como exemplo, cita-se mais uma vez os processos biológicos relacionado à degradação de drogas e à regulação negativa da apoptose por via extrínseca.

Além dos lncRNAs, nesse trabalho foi analisada a expressão de diferentes isoformas de um mesmo gene. Pouco se sabe sobre a função de isoformas específicas, apesar dos estudos de funções gênicas serem amplos. Ainda que a existência de isoformas de *splicing* alternativo não forneça suporte direto para sua funcionalidade ou distinção funcional (BHUIYAN et al., 2018), a modulação do *splicing* devido ao tratamento com cisplatina mostrou-se importante para a regulação dos mecanismos em que os genes modulados estão envolvidos.

Células humanas de carcinoma colorretal, HCT116, submetidas à hipóxia apresentaram uma alteração na expressão de isoformas de transcritos codificadores para isoformas não codificadoras. Essa alteração é consistente com o rápido declínio na síntese de proteínas observado em condições de hipóxia. A hipóxia levou ao aumento de retenção de íntrons em genes envolvidos na resposta ao dano de DNA e em genes envolvidos nas vias de reparo de DNA (MEMON et al., 2016).

Dlamini e colaboradores (DLAMINI et al., 2019), revisaram a contribuição do *splicing* alternativo na regulação da hipertensão induzida pela obesidade. Os autores concluíram que, uma vez que diferentes isoformas de *splicing* alternativo podem atuar de forma conflitante no metabolismo, a regulação do *splicing* ou o *splicing* anormal podem contribuir para o desenvolvimento de obesidade e de hipertensão relacionada a esta condição. Assim, variantes de *splicing* podem ser utilizadas como marcadores de diagnóstico e prognóstico, bem como podem ser alvos para o desenvolvimento de medicamentos.

O mecanismo de tumorigênese do carcinoma renal de células claras e aqueles envolvidos na metástase desta doença ainda não estão esclarecidos. Recentemente, os mecanismos de *splicing* alternativo foram avaliados para auxiliar na compreensão da evolução deste tipo de tumor. Meng e colaboradores (MENG et al., 2019), analisaram 537 amostras de RNA-Seq disponíveis no banco de dados TCGA e identificaram que a retenção de íntron nos oncogenes RHOT2 e TCIRG1 pelo fator de *splicing* DDX39B podem estar associadas à tumorigênese, metástase e mau prognóstico do carcinoma renal de células claras.

A análise ao nível de transcritos de rins de camundongos tratados com cisplatina mostrou que alguns genes tiveram apenas isoformas do grupo TPCP diferencialmente expressas. Isso significa que, apesar de diferencialmente expressos, esses mRNAs não serão traduzidos em proteínas, o que culmina na inexistência da função proteica. Esse foi o caso do

gene *Cpt1c*, que apresentou apenas isoforma classificada como íntron retido superexpressa. O gene *Cpt1c* ainda não tem papel elucidado na nefrotoxicidade induzida por cisplatina, entretanto, Maeda e colaboradores (MAEDA et al., 2014) relataram que este gene foi superexpresso em células tumorais resistentes à cisplatina. Este mesmo gene está relacionado à beta oxidação de ácidos graxo, de acordo com a atribuição de termos GO.

Alguns genes apresentaram tanto isoformas do grupo CP quanto do grupo TPCP diferencialmente expressas. Nesses casos, o nível de expressão geral do gene não reflete o nível de mRNAs aptos à tradução, ainda que a função proteica não seja completamente perdida, como é o caso do gene *Mdm2* anteriormente citado. A alteração das isoformas expressas ajuda a explicar a dissociação entre os níveis de expressão de RNA e proteína encontrada por alguns autores, como foi o caso de Späth e colaboradores (SPÄTH et al., 2019), responsáveis pela geração dos dados do Estudo 2. Outro fator que auxilia nessa compreensão é a existência da transcrição *readthrough* como será discutido a seguir.

A transcrição *readthrough* ocorre quando a transcrição continua após a região de término da transcrição. Estresse osmótico (VILBORG et al., 2015), câncer (GROSSO et al., 2015) e infecção viral (RUTKOWSKI et al., 2015) são condições que sabidamente levam a esse fenômeno. Pouco ainda se sabe sobre esse tipo de transcrição, todavia a importância dos transcritos gerados a partir da transcrição *readthrough* é destacada pela inserção deste tipo de transcrito na base de dados Ensembl em julho de 2019 (ENSEMBL, 2019). As anotações de genomas de camundongo e de humano passaram a conter dois novos tipos de genes e transcritos, são eles *readthrough* e códon de parada *readthrough* (*stop codon readthrough*). Genes *readthrough* são definidos como aqueles que possuem éxons pertencentes a dois ou mais loci além do locus que gerou o início da transcrição, já os do tipo códon de parada *readthrough* são aqueles em que a sequência de codificação contém um códon de parada traduzido e que a terminação ocorre em um códon de parada canônico mais a jusante (ENSEMBL, 2020).

A transcrição *readthrough* pode ser contínua o suficiente para adentrar ao gene à jusante daquele em que a transcrição foi iniciada, como foi observado por Grosso e colaboradores (GROSSO et al., 2015). Os autores identificaram que o gene antiapoptótico *BCL2* é frequentemente superexpresso em câncer renal de células claras e que o aumento de sua expressão está associado ao aumento da transcrição *readthrough* do gene à montante *KDSR*, demonstrando que a falha no término da transcrição interfere na expressão do gene à jusante. Tanto o aumento dos níveis do mRNA quanto da proteína BCL2 correlacionam-se com o aumento da transcrição *readthrough* de *KDSR*.

Em outros casos, entretanto, tem sido demonstrado que os transcritos gerados por transcrição *readthrough* sofrem *splicing* intergênico, dando origem a RNAs quiméricos (GROSSO et al., 2015; RUTKOWSKI et al., 2015). Nesse caso, os genes que sofrem transcrição *readthrough* e o gene à jusante, que formam o RNA quimérico, são induzidos transcricionalmente, porém não são traduzidos (RUTKOWSKI et al., 2015). A transcrição *readthrough* também pode ser interrompida na região intergênica, sem afetar o gene à jusante. Nesses casos, a região intergênica transcrita permanece ligada ao mRNA do gene que sofreu *readthrough*, como descrito por Vilborg e colaboradores (VILBORG et al., 2015) em condições de estresse osmótico. Os estresses osmótico e oxidativo e o choque térmico induzem transcrição *readthrough* e esses transcritos permanecem nucleares, aparentemente auxiliando na manutenção da integridade nuclear (VILBORG et al., 2017). Por serem transcritos de localização nuclear, não haverá tradução, levando a uma indução transcricional sem expressão proteica, como nos RNAs quiméricos supracitados.

No presente trabalho, a ocorrência de transcrição *readthrough* devido ao tratamento com cisplatina e na nefrotoxicidade está sendo mostrada pela primeira vez. O tratamento com cisplatina levou ao aumento de quase três vezes no número de DoGs gerados em relação aos animais não tratados. A expressão dos DoGs e seu tamanho mediano também se mostraram aumentados devido ao tratamento. Aparentemente, a transcrição *readthrough* faz parte da resposta celular ao estresse, conservada entre diferentes tipos de estresse e entre mamíferos. Uma hipótese é que a transcrição *readthrough* sirva para sequestrar Pol II levando a uma redução global da transcrição como parte da resposta adaptativa ao estresse (VILBORG et al., 2017).

De fato, a nefrotoxicidade causada por cisplatina tanto induz a transcrição *readthrough* quanto leva à repressão da transcrição que é refletida no grande número de transcritos subexpressos, que supera o número de superexpressos. Considerando que os transcritos gerados na transcrição *readthrough* não são traduzidos, como relatado por Vilborg (VILBORG et al., 2015) e Rutkowski (RUTKOWSKI et al., 2015), pode-se inferir que os processos biológicos enriquecidos nos genes que sofrem esse tipo de transcrição são afetados negativamente. Assim, a modulação nos padrões de *splicing* alternativo causada pela cisplatina é compatível com os processos biológicos relacionados ao processamento do mRNA enriquecidos dentre os genes que sofreram transcrição *readthrough*. Nesse sentido, também podemos inferir um prejuízo no transporte e localização de proteínas renais após o tratamento com cisplatina, devido ao enriquecimento destes processos biológicos. Destaca-se,

no entanto, a necessidade de avaliar mais profundamente o papel da transcrição *readthrough* na nefrotoxicidade induzida por cisplatina com estudos direcionados a esse fim.

Ainda não está claro o papel fisiológico da transcrição *readthrough* e dos DoGs. Apesar disso, esse fenômeno foi observado nos rins dos camundongos não tratados com cisplatina, que é refletido nos casos de DoGs subexpressos após o tratamento. Evidências sugerem que, em geral, os DoGs são regulados independentemente. Entretanto, é possível que uma regulação acentuada, seja positiva ou negativa, de um gene possa afetar os níveis de DoGs em direção semelhante (VILBORG et al., 2017). De fato, o tratamento com cisplatina induz uma forte subexpressão gênica que pode se relacionar com a subexpressão de DoGs na mesma condição. Essa observação é suportada pelos processos biológicos relacionados ao metabolismo basal que são enriquecidos tanto dentre os transcritos subexpressos quanto entre genes que sofrem transcrição *readthrough* e geram DoGs subexpressos.

6 CONCLUSÕES

No presente trabalho, avaliamos a modulação da expressão gênica na nefrotoxicidade induzida por cisplatina, com foco em isoformas de transcritos e genes não codificadores, bem como na avaliação de eventos de transcrição *readthrough*. Nossos resultados mostram que análise ao nível de transcritos recompõe processos biológicos conhecidos na nefrotoxicidade induzida por cisplatina quando analisados ao nível de genes. Porém, a mudança de expressão de isoformas de transcritos codificadores para isoformas não codificadoras, ou a existência de ambas, participam da modulação dos processos biológicos ao modificar os níveis de mRNAs aptos à tradução, como mostrado para o gene *Mdm2*, que apresentou cinco diferentes isoformas de transcritos superexpressos sendo apenas uma isoforma codificadora de proteína.

Mostramos que ao avaliar os processos biológicos é fundamental considerar o contexto que está sendo estudado. O enriquecimento do processo de regulação negativa da apoptose via extrínseca ocorreu devido à superexpressão de transcritos sabidamente envolvidos em outros mecanismos moleculares, sendo a apoptose por via extrínseca um desfecho comum ao tratamento com cisplatina. Já o enriquecimento do processo de catabolismo de drogas, se deve à subexpressão de genes da subfamília CYP2J. Identificamos que enzimas desta família não metabolizam cisplatina e seu papel na nefrotoxicidade está relacionado à regulação de transporte renal de íons e morte celular.

Nossas análises mostram que, ao integrar informações sobre os genes não codificadores de proteínas, a mecanismos de nefrotoxicidade conhecidos, foi possível acrescentar passos importantes aos mecanismos de internalização celular da cisplatina, regulação da apoptose e da inflamação, através da expressão dos lncRNAs *Airn*, *lincRNA-p21* e *Gm26917*, *Mirt1* e *Gn26669*, respectivamente.

Apresentamos, pela primeira vez, a existência de transcrição *readthrough* induzida pela cisplatina na nefrotoxicidade e que este tipo de transcrição pode afetar o processamento de mRNAs, transporte e localização de proteínas, além do metabolismo basal celular.

Concluimos que a avaliação integrada de expressão de isoformas de transcritos codificadores e não codificadores gerados por genes codificadores, bem como a expressão de genes não codificadores e de eventos de transcrição *readthrough* permitiu agregar passos a mecanismos de toxicidade previamente descritos, revelando novos possíveis alvos de estudos de biomarcadores e de intervenção terapêutica na prevenção da nefrotoxicidade induzida por cisplatina.

7 PERSPECTIVAS

O prosseguimento do trabalho aqui apresentado será conduzido com o intuito de aprofundamento e validação dos achados.

Tendo em vista que a mudança de isoformas de transcritos se mostrou importante para a nefrotoxicidade induzida por cisplatina, será realizada a montagem *de novo* do transcriptoma de rins a fim de identificar novas isoformas não anotadas. A validação, por RT-qPCR, de isoformas de transcritos será realizada em experimentos *in vitro* de células renais humanas tratadas com cisplatina. Além disso, investigaremos se as diferentes isoformas de transcritos codificadores estão presentes no proteoma de células renais ou nos rins de animais tratados com cisplatina. Assim como neste trabalho, esse novo passo será realizado à medida que dados públicos estiverem disponíveis.

Em relação à transcrição *readthrough*, experimentos *in vitro* serão planejados para validação dos DoGs, bem como para verificar se há enriquecimento dos mesmos nas frações nuclear e ou citoplasmática. A montagem *de novo* do transcriptoma também será utilizada para identificação de DoGs.

REFERÊNCIAS

- ADDINGTON, A. M. et al. A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. **Molecular Psychiatry**, v. 16, n. 3, p. 238–239, 18 mar. 2011.
- AIRES, M. de M. **Fisiologia**. 5. ed. Rio de Janeiro: Guanabara Koogan, 2018.
- ALBERTS, B. et al. **Biologia Molecular da Célula**. 6. ed. Porto Alegre: Artmed, 2017.
- ALEXA, A.; RAHNENFÜHRER, J. **topGO: enrichment analysis for Gene Ontology**, 2019.
- ALEXA, A.; RAHNENFÜHRER, J.; LENGAUER, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. **Bioinformatics**, v. 22, n. 13, p. 1600–1607, 1 jul. 2006.
- ANDERGASSEN, D. et al. The Airn lncRNA does not require any DNA elements within its locus to silence distant imprinted genes. **PLOS Genetics**, v. 15, n. 7, p. e1008268, 22 jul. 2019.
- ASHBURNER, M. et al. Gene ontology: tool for the unification of biology. **Nature Genetics**, v. 25, n. 1, p. 25–29, maio 2000.
- ATHIRAKUL, K. et al. Increased blood pressure in mice lacking cytochrome P450 2J5. **FASEB Journal**, v. 22, n. 12, p. 4096–4108, dez. 2008.
- BABRAHAM BIOINFORMATICS. **FastQC A Quality Control tool for High Throughput Sequence Data**. Disponível em: <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>>.
- BAI, W. et al. Long non-coding RNA NR_045623 and NR_028291 involved in benzene hematotoxicity in occupationally benzene-exposed workers. **Experimental and Molecular Pathology**, v. 96, n. 3, p. 354–360, jun. 2014.
- BARILE, F. A. **Clinical toxicology: principles and mechanisms**. 2. ed. New York: Informa Healthcare, 2010.
- BARNES, P. J.; KARIN, M. Nuclear factor- κ B - A pivotal transcription factor in chronic inflammatory diseases. **New England Journal of Medicine**, v. 336, n. 15, p. 1066–1071, 10 abr. 1997.
- BHATT, K. et al. MicroRNA-34a is induced via p53 during cisplatin nephrotoxicity and contributes to cell survival. **Molecular Medicine**, v. 16, n. 9–10, p. 409–416, set. 2010.
- BHUIYAN, S. A. et al. Systematic evaluation of isoform function in literature reports of alternative splicing. **BMC Genomics**, v. 19, n. 1, p. 637, 28 dez. 2018.

BLUME, C. J. et al. p53-dependent non-coding RNA networks in chronic lymphocytic leukemia. **Leukemia**, v. 29, n. 10, p. 2015–2023, 14 out. 2015. Disponível em: <<http://www.nature.com/articles/leu2015119>>. Acesso em: 5 set. 2019

BRANDENBURGER, T. et al. Noncoding RNAs in acute kidney injury. **Kidney International** Elsevier, v. 94, n. 5, p. 870-881, 2018.

BRAUNSCHWEIG, U. et al. Widespread intron retention in mammals functionally tunes transcriptomes. **Genome Research**, v. 24, n. 11, p. 1774–1786, 1 nov. 2014.

BRAY, N. L. et al. Near-optimal probabilistic RNA-seq quantification. **Nature Biotechnology**, v. 34, n. 5, p. 525–527, 2016.

BUNEL, V. et al. Early detection of acute cisplatin nephrotoxicity: Interest of urinary monitoring of proximal tubular biomarkers. **Clinical Kidney Journal**, v. 10, n. 5, p. 639–647, 1 out. 2017.

CAMPOS, M. A. A. et al. In vitro evaluation of biomarkers of nephrotoxicity through gene expression using gentamicin. **Journal of Biochemical and Molecular Toxicology**, v. 32, n. 9, p. e22189, 2018.

CARLSON, M. **org.Mm.eg.db: Genome wide annotation for Mouse**, 2019.

CARRIERI, C. et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. **Nature**, v. 491, n. 7424, p. 454–457, 15 nov. 2012.

CARTAULT, F. et al. A new XPC gene splicing mutation has lead to the highest worldwide prevalence of xeroderma pigmentosum in black Mahori patients. **DNA Repair**, v. 10, n. 6, p. 577–585, 10 jun. 2011.

CASTELLANOS-RUBIO, A. et al. A long noncoding RNA associated with susceptibility to celiac disease. **Science**, v. 352, n. 6281, p. 91–95, 1 abr. 2016.

CESANA, M. et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. **Cell**, v. 147, n. 2, p. 358–369, 14 out. 2011.

CHANDLER, D. S. et al. Genotoxic Stress Induces Coordinately Regulated Alternative Splicing of the p53 Modulators MDM2 and MDM4. **Cancer Research**, v. 66, n. 19, p. 9502–9508, 1 out. 2006.

CHEN, J.-K.; CAPDEVILA, J.; HARRIS, R. C. Cytochrome P450 Epoxygenase Metabolism of Arachidonic Acid Inhibits Apoptosis. **Molecular and Cellular Biology**, v. 21, n. 18, p. 6322–6331, 15 set. 2001.

CHEN, Y. et al. Down regulation of lincRNA-p21 contributes to gastric cancer development through Hippo-independent activation of YAP. **Oncotarget**, v. 8, n. 38, p. 63813–63824, 8 set. 2017. Disponível em: <<http://www.oncotarget.com/fulltext/19130>>. Acesso em: 5 set. 2019

CHEN, Y. et al. Long non-coding RNA NEAT1 plays an important role in sepsis-induced acute kidney injury by targeting miR-204 and modulating the NF-κB pathway. **International Immunopharmacology**, v. 59, p. 252–260, 1 jun. 2018a.

CHEN, Z. et al. Forkhead Box M1 Transcriptionally Regulates the Expression of Long Noncoding RNAs Snhg8 and Gm26917 to Promote Proliferation and Survival of Muscle Satellite Cells. **Stem Cells**, v. 36, n. 7, p. 1097–1108, jul. 2018b.

CHO, W. Y. et al. Urinary tissue inhibitor of metalloproteinase-2 and insulin-like growth factor-binding protein 7 as biomarkers of patients with established acute kidney injury. **The Korean Journal of Internal Medicine**, 23 dez. 2019.

CIARIMBOLI, G. et al. Organic cation transporter 2 mediates cisplatin-induced oto- and nephrotoxicity and is a target for protective interventions. **American Journal of Pathology**, v. 176, n. 3, p. 1169–1180, 2010.

CIARIMBOLI, G. Membrane transporters as mediators of cisplatin side-effects. **Anticancer research**, v. 34, n. 1, p. 547–50, 1 jan. 2014.

CONSTANZO, L. S. Fisiologia renal. In: **Fisiologia**. 6. ed. Rio de Janeiro: Elsevier, 2018.

DASARI, S.; TCHOUNWOU, P. B. Cisplatin in cancer therapy: Molecular mechanisms of action. **European Journal of Pharmacology**, v. 740, p. 364–378, 2014.

DAVIES, M. S.; BERNERS-PRICE, S. J.; HAMBLEY, T. W. Rates of platination of -AG- and -GA-containing double-stranded oligonucleotides: effect of chloride concentration. **Journal of Inorganic Biochemistry**, v. 79, n. 1–4, p. 167–172, 30 abr. 2000.

DERRIEN, T. et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. **Genome Research**, v. 22, n. 9, p. 1775–1789, set. 2012.

DING, G. et al. LincRNA-p21 inhibits invasion and metastasis of hepatocellular carcinoma through miR-9/E-cadherin cascade signaling pathway molecular mechanism. **OncoTargets and Therapy**, v. 10, p. 3241–3247, jun. 2017.

DJEBALI, S. et al. Landscape of transcription in human cells. **Nature**, v. 489, n. 7414, p. 101–108, 6 set. 2012.

DLAMINI, Z. et al. Regulation of alternative splicing in obesity-induced hypertension. **Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy**, v. 12, p. 1597–1615, 28 ago. 2019.

DOBIN, A. et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**, v. 29, n. 1, p. 15–21, 2012.

DVINGE, H.; BRADLEY, R. K. Widespread intron retention diversifies most cancer transcriptomes. **Genome Medicine**, v. 7, n. 1, p. 45, 15 dez. 2015.

EBRALIDZE, A. K. et al. PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. **Genes and Development**, v. 22, n. 15, p. 2085–2092, 1 ago. 2008.

ENSEMBL. **Gene and transcript types**. Disponível em: <http://vega.archive.ensembl.org/info/about/gene_and_transcript_types.html>. Acesso em: 26 ago. 2019.

ENSEMBL. **Ensembl 97 and Ensembl Genomes 44 have been released!** Disponível em: <<http://www.ensembl.info/2019/07/03/ensembl-97-and-ensembl-genomes-43-have-been-released/>>. Acesso em: 19 set. 2019.

ENSEMBL. **Biotypes**. Disponível em: <<https://www.ensembl.org/info/genome/genebuild/biotypes.html>>. Acesso em: 11 maio. 2020.

ESWARAN, J. et al. RNA sequencing of cancer reveals novel splicing alterations. **Scientific Reports**, v. 3, n. 1, p. 1689, 22 dez. 2013.

EWELS, P. et al. MultiQC: Summarize analysis results for multiple tools and samples in a single report. **Bioinformatics**, v. 32, n. 19, p. 3047–3048, 1 out. 2016.

FANG, S. et al. NONCODEV5: A comprehensive annotation database for long non-coding RNAs. **Nucleic Acids Research**, v. 46, n. D1, p. D308–D314, 2018.

FENG, Y. et al. Dysregulation of lncRNAs GM5524 and GM15645 involved in high-glucose-induced podocyte apoptosis and autophagy in diabetic nephropathy. **Molecular Medicine Reports**, v. 18, n. 4, p. 3657–3664, 21 ago. 2018.

FICHTINGER-SCHEPMAN, A. M. J. et al. Adducts of the Antitumor Drug cis-Diamminedichloroplatinum(II) with DNA: Formation, Identification, and Quantitation. **Biochemistry**, v. 24, n. 3, p. 707–713, jan. 1985.

FILIPSKI, K. K. et al. Contribution of organic cation transporter 2 (OCT2) to cisplatin-induced nephrotoxicity. **Clinical Pharmacology and Therapeutics**, v. 86, n. 4, p. 396–402, 2009.

FORTRIE, G.; DE GEUS, H. R. H.; BETJES, M. G. H. The aftermath of acute kidney injury: A narrative review of long-term mortality and renal function. **Critical Care**, v. 23, n. 1, p. 1–11, 2019.

FRANKISH, A. et al. GENCODE reference annotation for the human and mouse genomes. **Nucleic Acids Research**, v. 47, n. D1, p. D766–D773, 8 jan. 2019.

FUKASAWA, H. et al. Anti-Cancer Agent-Induced Nephrotoxicity. **Anti-Cancer Agents in Medicinal Chemistry**, v. 14, p. 921–927, 2014.

GALGAMUWA, R. et al. Dichloroacetate Prevents Cisplatin-Induced Nephrotoxicity without Compromising Cisplatin Anticancer Properties. **Journal of the American Society of Nephrology: JASN**, v. 27, n. 11, p. 3331–3344, 1 nov. 2016.

GENCODE. **Mouse: Statistics about the current GENCODE Release (version M20)**. Disponível em: <<https://www.genecodegenes.org/mouse/stats.html>>. Acesso em: 10 set. 2019.

GENCODE. **Human: Statistics about the current GENCODE Release (version 32)**. Disponível em: <<https://www.genecodegenes.org/human/stats.html>>. Acesso em: 10 set. 2019.

- GONG, C.; MAQUAT, L. E. LncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. **Nature**, v. 470, n. 7333, p. 284–290, 10 fev. 2011.
- GRAVES, J. P. et al. Characterization of four new mouse cytochrome P450 enzymes of the CYP2J subfamily. **Drug Metabolism and Disposition**, v. 41, n. 4, p. 763–773, abr. 2013.
- GRAVES, J. P. et al. Quantitative polymerase chain reaction analysis of the mouse Cyp2j subfamily: Tissue distribution and regulation. **Drug Metabolism and Disposition**, v. 43, n. 8, p. 1169–1180, 1 ago. 2015.
- GROSSI, M. F. et al. In vitro study of potential nephrotoxicity biomarkers through gene expression using Amphotericin B. **Journal of Toxicology and Pharmacology**, v. 1, n. 2, p. 1:008, 2017.
- GROSSO, A. R. et al. Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. **eLife**, v. 4, n. NOVEMBER2015, 17 nov. 2015.
- GÜÇLÜ, A. et al. MicroRNA-125b as a new potential biomarker on diagnosis of renal ischemia–reperfusion injury. **Journal of Surgical Research**, v. 207, p. 541–248, 1 jan. 2017.
- GUO, Y. et al. MicroRNA-709 Mediates Acute Tubular Injury through Effects on Mitochondrial Function. **Journal of the American Society of Nephrology**, v. 29, n. 2, p. 449–461, 1 fev. 2018.
- GUPTA, R. A. et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. **Nature**, v. 464, n. 7291, p. 1071–1076, 15 abr. 2010.
- HALL, J. E. **Tratado de fisiologia médica**. 13. ed. Rio de Janeiro: Elsevier, 2017.
- HAN, W. K. et al. Kidney Injury Molecule-1 (KIM-1): A novel biomarker for human renal proximal tubule injury. **Kidney International**, v. 62, n. 1, p. 237–244, 1 jul. 2002.
- HANSEN, T. B. et al. Natural RNA circles function as efficient microRNA sponges. **Nature**, v. 495, n. 7441, p. 384–388, 21 mar. 2013.
- HARTMANN, J. T. et al. Platinum organ toxicity and possible prevention in patients with testicular cancer. **International Journal of Cancer**, v. 83, n. 6, p. 866–869, 1999.
- HÄSLER, R. et al. Alterations of pre-mRNA splicing in human inflammatory bowel disease. **European Journal of Cell Biology**, v. 90, n. 6–7, p. 603–611, 1 jun. 2011.
- HASTINGS, M. L. et al. Expression of the thyroid hormone receptor gene, *erbA*, in B lymphocytes: Alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. **Nucleic Acids Research**, v. 25, n. 21, p. 4296–4300, 1 nov. 1997.
- HAUCK, W. W.; DONNER, A. Wald's test as applied to hypotheses in logit analysis. **Journal of the American Statistical Association**, v. 72, n. 360, p. 851–853, 1977.

HE, C. et al. The Role of Long Intergenic Noncoding RNA p21 in Vascular Endothelial Cells. **DNA and Cell Biology**, v. 34, n. 11, p. 677–683, nov. 2015.

HE, R. Z.; LUO, D. X.; MO, Y. Y. Emerging roles of lncRNAs in the post-transcriptional regulation in cancer. **Genes & diseases**, v. 6, n. 1, p. 6–15, Mar. 2019.

HELLER, M. J. DNA Microarray Technology: Devices, Systems, and Applications. **Annual Review of Biomedical Engineering**, v. 4, n. 1, p. 129–153, ago. 2002.

HIGBY, D. J. et al. Diaminodichloroplatinum: A phase I study showing responses in testicular and other tumors. **Cancer**, v. 33, n. 5, p. 1219–1225, 1 maio 1974.

HUANG, W. et al. Long non-coding RNA PVT1 promote LPS-induced septic acute kidney injury by regulating TNF α and JNK/NF- κ B pathways in HK-2 cells. **International Immunopharmacology**, v. 47, p. 134–140, 1 jun. 2017.

HUARTE, M. et al. A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response. **Cell**, v. 142, n. 3, p. 409–419, ago. 2010.

ICHIMURA, T. et al. Kidney Injury Molecule-1 (KIM-1), a Putative Epithelial Cell Adhesion Molecule Containing a Novel Immunoglobulin Domain, Is Up-regulated in Renal Cells after Injury. **Journal of Biological Chemistry**, v. 273, n. 7, p. 4135–4142, 13 fev. 1998.

ICHIMURA, T. et al. Kidney injury molecule-1 is a phosphatidylserine receptor that confers a phagocytic phenotype on epithelial cells. **Journal of Clinical Investigation**, v. 118, n. 5, p. 1657–1668, 1 maio 2008.

ILLUMINA. **TruSeq® DNA Sample Preparation Guide**. [s.l.: s.n.]. Disponível em: <https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_truseq/truseqrna/truseq_rna_sampleprep_guide_15008136_a.pdf>. Acesso em: 5 ago. 2019

ISHIDA, S. et al. Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. **Proceedings of the National Academy of Sciences of the United States of America**, v. 99, n. 22, p. 14298–14302, 2002.

JIANG, M. et al. Regulation of PUMA- α by p53 in cisplatin-induced renal cell apoptosis. **Oncogene**, v. 25, n. 29, p. 4056–4066, 6 jul. 2006.

JOO, M. S. et al. miR-125b transcriptionally increased by Nrf2 inhibits AhR repressor, which protects kidney from cisplatin-induced injury. **Cell Death and Disease**, v. 4, n. 10, p. e899, out. 2013.

JUMAA, H.; NIELSEN, P. J. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. **The EMBO Journal**, v. 16, n. 16, p. 5077–5085, 15 ago. 1997.

JURICA, M. S.; MOORE, M. J. Pre-mRNA Splicing. **Molecular Cell**, v. 12, n. 1, p. 5–14, 1 jul. 2003.

KAMELGARN, M. et al. ALS mutations of FUS suppress protein translation and disrupt the regulation of nonsense-mediated decay. **Proceedings of the National Academy of Sciences**, v. 115, n. 51, p. E11904–E11913, 18 dez. 2018.

KANDURI, C. Long noncoding RNAs: Lessons from genomic imprinting. **Biochimica et Biophysica Acta - Gene Regulatory Mechanisms**, v. 1859, p. 102-111, jan. 2016.

KANS, J. Entrez Direct: E-utilities on the UNIX Command Line. 2013 Apr 23 [Updated 2017 Jul 10]. In: **Entrez Programming Utilities Help [Internet]**. Bethesda (Md): National Center for Biotechnology Information (US), 2010.

KAUSHAL, G. P. et al. Role and regulation of activation of caspases in cisplatin-induced injury to renal tubular epithelial cells. **Kidney International**, v. 60, n. 5, p. 1726–1736, nov. 2001.

KDIGO ACUTE KIDNEY INJURY WORK GROUP. KDIGO clinical practice guideline for acute kidney injury. **Kidney International Supplements**, v. 2, n. 1, p. 1–138, 2012.

KELLAND, L. The resurgence of platinum-based cancer chemotherapy. **Nature Reviews Cancer**, v. 7, n. 8, p. 573–584, 2007.

KELLUM, J. A.; RONCO, C.; BELLOMO, R. Acute kidney disease and the community. **The Lancet**, v. 387, n. 10032, p. 1974-1976, 2016.

KHAN, A. H. et al. Novel orally active epoxyeicosatrienoic acid (EET) analogs attenuate cisplatin nephrotoxicity. **FASEB Journal**, v. 27, n. 8, p. 2946–2956, 2013.

KHAN, A. H. et al. Epoxyeicosatrienoic acid analogue lowers blood pressure through vasodilation and sodium channel inhibition. **Clinical Science**, v. 127, n. 7, p. 463–474, 2014.

KOMERS, R. et al. Transcriptome-based analysis of kidney gene expression changes associated with diabetes in OVE26 mice, in the presence and absence of losartan treatment. **PLoS ONE**, v. 9, n. 5, p. e96987, 14 maio 2014.

KONG, Q.; QIU, M. Long noncoding RNA SNHG15 promotes human breast cancer proliferation, migration and invasion by sponging miR-211-3p. **Biochemical and Biophysical Research Communications**, v. 495, n. 2, p. 1594–1600, jan. 2018.

KREBS, J. E.; KILPATRICK, S. T.; GOLDSTEIN, E. S. **Lewin's Genes XI**. 11. ed. Burlington: Jones & Bartlett Learning, 2014.

KROETZ, D. L.; ZELDIN, D. C. Cytochrome P450 pathways of arachidonic acid metabolism. **Current Opinion in Lipidology**, v. 13, n. 3, p. 273-283, 2002.

KUBBUTAT, M. H. G.; JONES, S. N.; VOUSDEN, K. H. Regulation of p53 stability by Mdm2. **Nature**, v. 387, n. 6630, p. 299–303, 15 maio 1997.

LAROCCA, J. et al. The impact of first trimester phthalate and phenol exposure on IGF2/H19 genomic imprinting and birth outcomes. **Environmental Research**, v. 133, p. 396–406, ago. 2014.

LASSNIGG, A. et al. Minimal changes of serum creatinine predict prognosis in patients after cardiothoracic surgery: A prospective cohort study. **Journal of the American Society of Nephrology**, v. 15, n. 6, p. 1597–1605, jun. 2004.

LATOS, P. A. et al. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. **Science**, v. 338, n. 6113, p. 1469–1472, 14 dez. 2012.

LAURENT, G. St. et al. Functional annotation of the vlinc class of non-coding RNAs using systems biology approach. **Nucleic Acids Research**, v. 44, n. 7, p. 3233–3252, 20 abr. 2016.

LEINONEN, R.; SUGAWARA, H.; SHUMWAY, M. The Sequence Read Archive. **Nucleic Acids Research**, v. 39, n. Database, p. D19–D21, 1 jan. 2011.

LEVIN, A. et al. Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy. **The Lancet**, v. 390, n. 10105, p. 1888–1917, 2017.

LI, W.; NOTANI, D.; ROSENFELD, M. G. Enhancers as non-coding RNA transcription units: Recent insights and future perspectives. **Nature Reviews Genetics**, v. 17, n. 4, p. 207–223, 1 abr. 2016.

LI, X.; ZHOU, J.; HUANG, K. Inhibition of the lncRNA Mirt1 Attenuates Acute Myocardial Infarction by Suppressing NF- κ B Activation. **Cellular Physiology and Biochemistry**, v. 42, n. 3, p. 1153–1164, 2017.

LIAO, B. et al. Evaluation of a Diagnostic Test of Serum Neutrophil Gelatinase-Associated Lipocalin (NGAL) and Urine KIM-1 in Contrast-Induced Nephropathy (CIN). **Medical Science Monitor**, v. 25, p. 565–570, 19 jan. 2019.

LIAO, W. et al. MicroRNA-140-5p attenuated oxidative stress in Cisplatin induced acute kidney injury by activating Nrf2/ARE pathway through a Keap1-independent mechanism. **Experimental Cell Research**, v. 360, n. 2, p. 292–302, 15 nov. 2017.

LINDEBOOM, R. G. H.; SUPEK, F.; LEHNER, B. The rules and impact of nonsense-mediated mRNA decay in human cancers. **Nature Genetics**, v. 48, n. 10, p. 1112–1118, 12 out. 2016.

LIU, G.; MATTICK, J. S.; TAFT, R. J. A meta-analysis of the genomic and transcriptomic composition of complex life. **Cell Cycle**, v. 12, n. 13, p. 2061–2072, 1 jul. 2013.

LIU, Y. et al. Down-regulation of myocardial infarction associated transcript 1 improves myocardial ischemia-reperfusion injury in aged diabetic rats by inhibition of activation of NF- κ B signaling pathway. **Chemico-Biological Interactions**, v. 300, p. 111–122, fev. 2019a.

LIU, Z. et al. Toxicogenomics: A 2020 Vision. **Trends in Pharmacological Sciences**, v. 40, n. 2, p. 92–103, 1 fev. 2019b.

LOWE, R. et al. Transcriptomics technologies. **PLoS Computational Biology**, v. 13, n. 5, p. e1005457, 1 maio 2017.

LYKKE-ANDERSEN, S.; JENSEN, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. **Nature Reviews Molecular Cell Biology**, v. 16, n. 11, p. 665–677, 23 nov. 2015.

MA, J. et al. Molecular cloning, enzymatic characterization, developmental expression, and cellular localization of a mouse cytochrome P450 highly expressed in kidney. **Journal of Biological Chemistry**, v. 274, n. 25, p. 17777–17788, 18 jun. 1999.

MAEDA, O. et al. Alteration of gene expression and DNA methylation in drug-resistant gastric cancer. **Oncology Reports**, v. 31, n. 4, p. 1883–1890, 1 abr. 2014.

MAGISTRI, M. et al. Transcriptomics Profiling of Alzheimer's Disease Reveal Neurovascular Defects, Altered Amyloid- β Homeostasis, and Deregulated Expression of Long Noncoding RNAs. **Journal of Alzheimer's Disease**, v. 48, n. 3, p. 647–665, 1 out. 2015.

MANDIC, A. et al. Cisplatin induces endoplasmic reticulum stress and nucleus-independent apoptotic signaling. **Journal of Biological Chemistry**, v. 278, n. 11, p. 9100–9106, 14 mar. 2003.

MANOHAR, S.; LEUNG, N. Cisplatin nephrotoxicity: a review of the literature. **Journal of Nephrology**, v. 31, n. 1, p. 15–25, 5 fev. 2018.

MARRA, M. A.; HILLIER, L.; WATERSTON, R. H. Expressed sequence tags--ESTablishing bridges between genomes. **Trends in genetics : TIG**, v. 14, n. 1, p. 4–7, 1998.

MARTENS, J. A.; LAPRADE, L.; WINSTON, F. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. **Nature**, v. 429, n. 6991, p. 571–574, 3 jun. 2004.

MATLIN, A. J.; CLARK, F.; SMITH, C. W. J. Understanding alternative splicing: towards a cellular code. **Nature Reviews Molecular Cell Biology**, v. 6, n. 5, p. 386–398, maio 2005.

MCILWAIN, D. R. et al. Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 27, p. 12186–12191, 6 jul. 2010.

MEDGHALCHI, S. M. et al. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. **Human Molecular Genetics**, v. 10, n. 2, p. 99–105, 1 jan. 2001.

MEERSCH, M. et al. Urinary TIMP-2 and IGFBP7 as Early Biomarkers of Acute Kidney Injury and Renal Recovery following Cardiac Surgery. **PLoS ONE**, v. 9, n. 3, p. e93460, 27 mar. 2014.

MEISTER, G. **RNA Biology: an introduction**. 1. ed. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2011.

MEMON, D. et al. Hypoxia-driven splicing into noncoding isoforms regulates the DNA damage response. **npj Genomic Medicine**, v. 1, n. 1, p. 16020, 20 nov. 2016.

MENG, T. et al. Identification of Prognostic and Metastatic Alternative Splicing Signatures in Kidney Renal Clear Cell Carcinoma. **Frontiers in Bioengineering and Biotechnology**, v. 7, p. 270, 15 out. 2019.

MEZENCEV, R.; SUBRAMANIAM, R. The use of evidence from high-throughput screening and transcriptomic data in human health risk assessments. **Toxicology and Applied Pharmacology**, v. 380, p. 114706, 1 out. 2019.

- MILLER, R. P. et al. Mechanisms of cisplatin nephrotoxicity. **Toxins**, v. 2, n. 11, p. 2490–2518, 2010.
- MORRIS, K. V.; MATTICK, J. S. The rise of regulatory RNA. **Nature Reviews Genetics**, v. 15, n. 6, p. 423-437, 2014.
- NAKAMURA, T. et al. Disruption of multidrug and toxin extrusion MATE1 potentiates cisplatin-induced nephrotoxicity. **Biochemical Pharmacology**, v. 80, n. 11, p. 1762–1767, 1 dez. 2010.
- NASH, K.; HAFEEZ, A.; HOU, S. Hospital-acquired renal insufficiency. **American Journal of Kidney Diseases**, v. 39, n. 5, p. 930-936, 2002.
- NATIONAL RESEARCH COUNCIL. **Toxicity Testing in the 21st Century**. Washington, D.C.: National Academies Press, 2007.
- NCBI. **The Sequence Read Archive (SRA) Overview**. Disponível em: <<https://www.ncbi.nlm.nih.gov/sra/docs/>>. Acesso em: 19 set. 2017a.
- NCBI. **Tool: fastq-dump**. Disponível em: <https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=toolkit_doc&f=fastq-dump>.
- NCBI. Aspera Transfer Guide. In: **SRA Handbook [Internet]**. Bethesda (Md): National Center for Biotechnology Information (US), 2014.
- NUWAYSIR, E. F. et al. Microarrays and toxicology: The advent of toxicogenomics. **Molecular Carcinogenesis**, v. 24, n. 3, p. 153–159, 1999.
- PABLA, N. et al. ATR-Chk2 signaling in p53 activation and DNA damage response during cisplatin-induced apoptosis. **Journal of Biological Chemistry**, v. 283, n. 10, p. 6572–6583, 7 mar. 2008.
- PABLA, N. et al. The copper transporter Ctr1 contributes to cisplatin uptake by renal tubular cells during cisplatin nephrotoxicity. **American Journal of Physiology - Renal Physiology**, v. 296, n. 3, p. F505–F511, 2009.
- PABLA, N.; DONG, Z. Cisplatin nephrotoxicity: Mechanisms and renoprotective strategies. **Kidney International**, v. 73, n. 9, p. 994–1007, 2008.
- PACHTER LAB. **Manual kallisto**. Disponível em: <<https://pachterlab.github.io/kallisto/manual>>. Acesso em: 5 ago. 2019.
- PAZHAYATTIL, G. S.; SHIRALI, A. C. Drug-induced impairment of renal function. **International Journal of Nephrology and Renovascular Disease**, v. 7, p. 457–458, 12 dez. 2014.
- PERES, L. A. B.; CUNHA JÚNIOR, A. D. da. Nefrotoxicidade aguda da cisplatina: mecanismos moleculares. **Jornal Brasileiro de Nefrologia**, v. 35, n. 4, p. 332–340, 2013.
- PERKINS, N. D. Integrating cell-signalling pathways with NF- κ B and IKK function. **Nature Reviews Molecular Cell Biology**, v. 8, n. 1, p. 49–62, jan. 2007.

PIMENTEL, H. et al. Differential analysis of RNA-seq incorporating quantification uncertainty. **Nature Methods**, v. 14, n. 7, p. 687–690, 2017.

QIAN, W. et al. Mitochondrial density determines the cellular sensitivity to cisplatin-induced cell death. **American Journal of Physiology-Cell Physiology**, v. 289, n. 6, p. C1466–C1475, dez. 2005.

QUINTANILHA, J. C. F. et al. Involvement of cytochrome P450 in cisplatin treatment: implications for toxicity. **Cancer Chemotherapy and Pharmacology**, v. 80, n. 2, p. 223–233, 1 ago. 2017.

RAMESH, G.; REEVES, W. B. TNF- α mediates chemokine and cytokine expression and renal injury in cisplatin nephrotoxicity. **Journal of Clinical Investigation**, v. 110, n. 6, p. 835–842, 15 set. 2002.

RAO, M. S. et al. Comparison of RNA-Seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term rat toxicity studies. **Frontiers in Genetics**, v. 9, p. 636, 2019.

RECIO, L. et al. Differential expression of long noncoding RNAs in the livers of female B6C3F1 mice exposed to the carcinogen furan. **Toxicological Sciences**, v. 135, n. 2, p. 369–379, out. 2013.

REN, G. et al. Noncoding RNAs in acute kidney injury. **Journal of Cellular Physiology**, v. 234, n. 3, p. 2266–2276, 26 mar. 2019.

RILEY, T. et al. Transcriptional control of human p53-regulated genes. **Nature Reviews Molecular Cell Biology**, v. 9, n. 5, p. 402–412, maio 2008.

ROSENBERG, B. et al. Platinum compounds: A new class of potent antitumour agents. **Nature**, v. 222, n. 5191, p. 385–386, 1969.

ROSENBERG, B. Cisplatin: Its history and possible mechanisms of action. In: PRESTAYKO, A. W.; CROOKE, S. T.; CARTER, S. K. (Ed.). **Cisplatin: current status and new developments**. [s.l.] Academic Press, 1980. p. 9–20.

ROSENBERG, B.; VAN CAMP, L.; KRIGAS, T. Inhibition of cell division in *Escherichia coli* by electrolysis products from a platinum electrode. **Nature**, v. 205, n. 4972, p. 698–699, 1965.

RUTKOWSKI, A. J. et al. Widespread disruption of host transcription termination in HSV-1 infection. **Nature Communications**, v. 6, n. 1, p. 7126, 20 nov. 2015.

RYAN, K.; BAUER, D. L. V. Finishing touches: Post-translational modification of protein factors involved in mammalian pre-mRNA 3' end formation. **The International Journal of Biochemistry & Cell Biology**, v. 40, n. 11, p. 2384–2396, 1 jan. 2008.

SAFIRSTEIN, R.; DAYE, M.; GUTTENPLAN, J. B. Mutagenic activity and identification of excreted platinum in human and rat urine and rat plasma after administration of cisplatin. **Cancer Letters**, v. 18, n. 3, p. 329–338, 1 abr. 1983.

SCHENA, M. et al. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science**, v. 270, n. 5235, p. 467–470, 1995.

SCHRAGENHEIM, J. et al. EET enhances renal function in obese mice resulting in restoration of HO-1-Mfn1/2 signaling, and decrease in hypertension through inhibition of sodium chloride co-transporter. **Prostaglandins and Other Lipid Mediators**, v. 137, p. 30–39, 1 jul. 2018.

SEQC/MAQC-III CONSORTIUM. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. **Nature Biotechnology**, v. 32, n. 9, p. 903–914, 1 set. 2014.

SERVAIS, H. et al. Renal cell apoptosis induced by nephrotoxic drugs: Cellular and molecular mechanisms and potential approaches to modulation. **Apoptosis**, v. 13, n. 1, p. 11–32, jan. 2008.

SETH, R. et al. p53-dependent caspase-2 activation in mitochondrial release of apoptosis-inducing factor and its role in renal tubular epithelial cell injury. **Journal of Biological Chemistry**, v. 280, n. 35, p. 31230–31239, 2 set. 2005.

SHARMA, A. et al. Calcium-mediated histone modifications regulate alternative splicing in cardiomyocytes. **Proceedings of the National Academy of Sciences**, v. 111, n. 46, p. E4920–E4928, 18 nov. 2014.

SHIRAKI, T. et al. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 26, p. 15776–15781, 23 dez. 2003.

SILVA, S. C. T. et al. In vitro study of putative genomic biomarkers of nephrotoxicity through differential gene expression using gentamicin. **Toxicology Mechanisms and Methods**, v. 27, n. 6, p. 435–441, 24 jul. 2017.

SINGH, R. K. et al. Conserved sequences in the final intron of MDM2 are essential for the regulation of alternative splicing of MDM2 in response to stress. **Experimental Cell Research**, v. 315, n. 19, p. 3419–3432, 15 nov. 2009.

SPÄTH, M. R. et al. The proteome microenvironment determines the protective effect of preconditioning in cisplatin-induced acute kidney injury. **Kidney International**, v. 95, n. 2, p. 333–349, 1 fev. 2019.

STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. **Nature Reviews Genetics**, v. 20, p. 631–656, 2019.

SUN, H. et al. RNA sequencing revealing the role of AMP-activated protein kinase signaling in mice myocardial ischemia reperfusion injury. **Gene**, v. 703, p. 91–101, jun. 2019.

SUPAVEKIN, S. et al. Differential gene expression following early renal ischemia/reperfusion. **Kidney International**, v. 63, n. 5, p. 1714–1724, 1 maio 2003.

SUREAU, A. et al. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. **The EMBO Journal**, v. 20, n. 7, p. 1785–1796, 2 abr. 2001.

SUTER, L.; BABISS, L. E.; WHEELDON, E. B. Toxicogenomics in Predictive Toxicology in Drug Development Review. **Chemistry & Biology**, v. 11, p. 161–171, 2004.

TABATA, T. et al. Interleukin-24 is a novel diagnostic biomarker for the severity of acute kidney injury. **Medical Molecular Morphology**, 2019.

TANGE, O. GNU Parallel: the command-line power tool. **The USENIX Magazine**, v. 36, n. 1, p. 42–47, fev. 2011.

TANI, H. et al. Long Non-Coding RNAs as Surrogate Indicators for Chemical Stress Responses in Human-Induced Pluripotent Stem Cells. **PLoS ONE**, v. 9, n. 8, p. e106282, 29 ago. 2014.

TANI, H. et al. Short-lived long non-coding RNAs as surrogate indicators for chemical exposure and LINC00152 and MALAT1 modulate their neighboring genes. **PLoS ONE**, v. 12, n. 7, p. e0181628, 18 jul. 2017.

TANI, H.; TORIMURA, M. Identification of short-lived long non-coding RNAs as surrogate indicators for chemical stress response. **Biochemical and Biophysical Research Communications**, v. 439, n. 4, p. 547–551, out. 2013.

THE GENE ONTOLOGY CONSORTIUM. The Gene Ontology Resource: 20 years and still GOing strong. **Nucleic Acids Research**, v. 47, n. D1, p. D330–D338, 8 jan. 2019.

TONG, J. et al. SNHG15: a promising cancer-related long noncoding RNA. **Cancer Management and Research**, v. Volume 11, p. 5961–5969, jul. 2019.

VAIDYA, V. S. et al. Kidney injury molecule-1 outperforms traditional biomarkers of kidney injury in preclinical biomarker qualification studies. **Nature Biotechnology**, v. 28, n. 5, p. 478–485, 2010.

VELCULESCU, V. E. et al. Serial analysis of gene expression. **Science**, v. 270, n. 5235, p. 484–487, 1995.

VELCULESCU, V. E. et al. Characterization of the yeast transcriptome. **Cell**, v. 88, n. 2, p. 243–251, 24 jan. 1997.

VILBORG, A. et al. Widespread Inducible Transcription Downstream of Human Genes. **Molecular Cell**, v. 59, n. 3, p. 449–461, 6 ago. 2015.

VILBORG, A. et al. Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. **Proceedings of the National Academy of Sciences**, v. 114, n. 40, p. E8362–E8371, 3 out. 2017.

VILBORG, A.; STEITZ, J. A. Readthrough transcription: How are DoGs made and what do they do? **RNA Biology**, v. 14, n. 5, p. 632–636, 2017.

WANG, C. et al. The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. **Nature Biotechnology**, v. 32, n. 9, p. 926–932, 24 set. 2014a.

WANG, C. et al. The Proangiogenic Roles of Long Non- Coding RNAs Revealed by RNA-Sequencing Following Oxygen-Glucose Deprivation/Re- Oxygenation. **Cellular Physiology and Biochemistry**, v. 52, n. 4, p. 708–727, 29 mar. 2019.

WANG, E. J. et al. Validation of putative genomic biomarkers of nephrotoxicity in rats. **Toxicology**, v. 246, n. 2–3, p. 91–100, 2008a.

WANG, E. T. et al. Alternative isoform regulation in human tissue transcriptomes. **Nature**, v. 456, n. 7221, p. 470–476, 27 nov. 2008b.

WANG, G. et al. LincRNA-p21 enhances the sensitivity of radiotherapy for human colorectal cancer by targeting the Wnt/ β -catenin signaling pathway. **Oncology Reports**, v. 31, n. 4, p. 1839–1845, abr. 2014b.

WANG, K. C. et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. **Nature**, v. 472, n. 7341, p. 120–126, 7 abr. 2011.

WANG, X. et al. LincRNA-p21 suppresses development of human prostate cancer through inhibition of PKM2. **Cell Proliferation**, v. 50, n. 6, p. e12395, dez. 2017.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: A revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, n. 1, p. 57–63, jan. 2009.

WEISS, R. et al. Acute kidney injury: a frequently underestimated problem in perioperative medicine. **Deutsches Ärzteblatt International**, v. 116, n. 49, p. 833–842, 6 dez. 2019.

WIESEL, Y.; SABATH, N.; SHALGI, R. DoGFinder: A software for the discovery and quantification of readthrough transcripts from RNA-seq. **BMC Genomics**, v. 19, n. 1, 2018.

WILLINGHAM, A. T. et al. Molecular biology: A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. **Science**, v. 309, n. 5740, p. 1570–1573, 2 set. 2005.

WU, G. et al. LincRNA-p21 regulates neointima formation, vascular smooth muscle cell proliferation, apoptosis, and atherosclerosis by enhancing p53 activity. **Circulation**, v. 130, n. 17, p. 1452–1465, 21 out. 2014.

XIAO, H. et al. RNA-Seq analysis of potential lncRNAs and genes for the anti-renal fibrotic effect of norcantharidin. **Journal of Cellular Biochemistry**, v. 120, n. 10, p. 17354–17367, 19 out. 2019.

XU, X. et al. LincRNA-p21 inhibits cell viability and promotes cell apoptosis in parkinson's disease through activating α -Synuclein Expression. **BioMed Research International**, v. 2018, 25 dez. 2018.

YONEYAMA, F. et al. Novel Urinary Biomarkers for Acute Kidney Injury and Prediction of Clinical Outcomes After Pediatric Cardiac Surgery. **Pediatric Cardiology**, 2019.

YONEZAWA, A. et al. Association between tubular toxicity of cisplatin and expression of organic cation transporter rOCT2 (Slc22a2) in the rat. **Biochemical Pharmacology**, v. 70, n. 12, p. 1823–1831, dez. 2005.

- YONEZAWA, A. et al. Cisplatin and oxaliplatin, but not carboplatin and nedaplatin, are substrates for human organic cation transporters (SLC22A1-3 and multidrug and toxin extrusion family). **Journal of Pharmacology and Experimental Therapeutics**, v. 319, n. 2, p. 879–886, 2006.
- YU, F. et al. LincRNA-p21 Inhibits the Wnt/ β -Catenin Pathway in Activated Hepatic Stellate Cells via Sponging MicroRNA-17-5p. **Cellular Physiology and Biochemistry**, v. 41, n. 5, p. 1970–1980, 2017.
- YU, T. M. et al. RANTES mediates kidney ischemia reperfusion injury through a possible role of HIF-1 α and LncRNA PRINS. **Scientific Reports**, v. 6, n. 18424, 4 jan. 2016.
- ZANGRANDO, J. et al. Identification of candidate long non-coding RNAs in response to myocardial infarction. **BMC Genomics**, v. 15, n. 1, p. 460, 2014.
- ZHANG, C. et al. Bioinformatics Tools for RNA-seq Gene and Isoform Quantification. **Journal of Next Generation Sequencing & Applications**, v. 03, n. 03, 2017.
- ZHANG, J. et al. Isolation and Characterization of a Novel Noncoding RNA from Nickel-Induced Lung Cancer. **Biological Trace Element Research**, v. 150, n. 1–3, p. 258–263, 5 dez. 2012.
- ZHANG, J. et al. LincRNA-p21 sponges miR-18b to promote the progression of diabetic nephropathy. **American journal of translational research**, v. 10, n. 5, p. 1481–1489, 2018a.
- ZHANG, J. et al. Identification and analysis of a key long non-coding RNAs (lncRNAs)-associated module reveal functional lncRNAs in cardiac hypertrophy. **Journal of Cellular and Molecular Medicine**, v. 22, n. 2, p. 892–903, 20 nov. 2018b.
- ZHANG, Y. et al. LincRNA-p21 leads to G1 arrest by p53 pathway in esophageal squamous cell carcinoma. **Cancer Management and Research**, v. Volume 11, p. 6201–6214, jul. 2019.
- ZHENG, J. et al. LincRNA-p21 inhibits hepatic stellate cell activation and liver fibrogenesis via p21. **FEBS Journal**, v. 282, n. 24, p. 4810–4821, dez. 2015.
- ZHOU, Q. et al. LincRNA-p21 suppresses glutamine catabolism and bladder cancer cell growth through inhibiting glutaminase expression. **Bioscience Reports**, v. 39, n. 4, p. BSR20182372, 30 abr. 2019.
- ZUK, A.; BONVENTRE, J. V. Acute kidney injury. **Annual Review of Medicine**, v. 67, p. 293–307, 14 jan. 2016.

APÊNDICE A – LISTA DE FÁRMACOS NEFROTÓXICOS

Acetaminophen	Diphenhydramine	Naftidrofuryl
Acetylsalicylic Acid	Doxepin	Nedaplatin
Aciclovir	Doxylamine	Neomycin
Adefovir	Erythromycin	Nitrosurea
Allopurinol	Fenacetin	Okt3
Amikacin	FK506	Omeprazole
Aminoglycoside	Folic Acid	Oxaliplatin
Amphotericin	Foscarnet	Pamidronate
Amphotericin B	Furosemide	Pantoprazole
Antifungal	Ganciclovir	Paracetamol
Antiparasitic	Gemcitabine	Penicillamine
Aristolochic Acid	Gentamicin	Penicillins
ASA (Acetylsalicylic Acid)	Gold Salts	Pentamidine
Atazanavir	Heavy Metal	Phenytoin
Azathioprine	Hydralazine	Propylthiouracil
Beta Lactam	Ifosfamide	Pyrazinamide
Beta-lactam	Il-2	Quinine
Bunamiodil	Immunoglobulin	Quinolone
Captopril	Indinavir	Rifampin
Carboplatin	Interferon Alpha	Ritonavir
Cephalosporins	Iopanoic Acid	Sirolimus
Cidofovir	Iothalamate	Statins
Cimetidine	Isoniazid	Streptomycin
Ciprofloxacin	Lansoprazole	Sucrose
Cisplatin	Lithium	Sulfonamide
Clopidogrel	Mannitol	Sulphadiazine
Cotrimoxazole	Mesalazine	
Cyclophosphamide	Methotrexate	
Cyclosporine	Methoxyflurane	
Dextran	Minocycline	
Diatrizoate	Mitomycin	

APÊNDICE B – TRANSCRITOS SUPEREXPRESSOS EM COMUM AO ESTUDO 1

E ESTUDO 2

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
1200007C13Rik	lincRNA	ENSMUST00000143649	lincRNA	2,959	0,00991	2,025	0,00005
1700007K13Rik	protein_coding	ENSMUST00000086370	protein_coding	5,058	0,00027	4,531	0,00002
1700072G22Rik	lincRNA	ENSMUST00000189460	lincRNA	4,174	0,00257	4,922	0,00002
1700110K17Rik	lincRNA	ENSMUST00000156511	antisense	6,326	0,00029	1,201	0,00064
2010003K11Rik	protein_coding	ENSMUST00000048482	protein_coding	2,302	0,00015	1,328	0,00068
2200002D01Rik	protein_coding	ENSMUST00000032808	protein_coding	1,621	0,00036	1,830	0,00000
2200002D01Rik	protein_coding	ENSMUST00000151339	retained_intron	1,553	0,00044	2,427	0,00000
2810459M11Rik	protein_coding	ENSMUST00000165824	protein_coding	1,719	0,00570	1,606	0,00046
3300005D01Rik	lincRNA	ENSMUST00000181484	lincRNA	2,271	0,00860	3,381	0,00030
4833422C13Rik	lincRNA	ENSMUST00000161491	antisense	1,229	0,00125	1,580	0,00011
4833428L15Rik	lincRNA	ENSMUST00000181230	antisense	3,377	0,00479	2,319	0,00016
5031415H12Rik	lincRNA	ENSMUST00000181546	antisense	1,776	0,00110	2,585	0,00014
5031425E22Rik	lincRNA	ENSMUST00000185943	lincRNA	1,973	0,00254	1,822	0,00038
5033403H07Rik	lincRNA	ENSMUST00000198580	lincRNA	4,159	0,00140	2,912	0,00325
5033403H07Rik	lincRNA	ENSMUST00000196858	retained_intron	3,935	0,00239	3,490	0,00062
9130230L23Rik	protein_coding	ENSMUST00000201351	protein_coding	3,446	0,00657	2,294	0,00261
9130230L23Rik	protein_coding	ENSMUST00000202278	retained_intron	1,825	0,00563	1,960	0,00001
9130230L23Rik	protein_coding	ENSMUST00000201562	retained_intron	1,701	0,00874	1,470	0,00005
A930001C03Rik	lincRNA	ENSMUST00000133871	processed_transcript	1,791	0,00170	2,002	0,00000
Aaas	protein_coding	ENSMUST00000041208	protein_coding	1,841	0,00005	1,347	0,00000
Aaas	protein_coding	ENSMUST00000230812	retained_intron	1,820	0,00049	1,966	0,00001
Abcb1b	protein_coding	ENSMUST00000009058	protein_coding	1,480	0,00006	2,594	0,00000
AC140316.1	pseudogene	ENSMUST00000235119	processed_pseudogene	3,657	0,00927	2,130	0,00024
AC167363.2	TEC	ENSMUST00000233069	TEC	4,412	0,00101	4,449	0,00001
Acot1	protein_coding	ENSMUST00000222591	retained_intron	1,429	0,00268	1,387	0,00047
Acot1	protein_coding	ENSMUST00000168120	protein_coding	1,330	0,00013	1,329	0,00464
Acot2	protein_coding	ENSMUST00000021649	protein_coding	1,079	0,00007	1,265	0,00022
Adm	protein_coding	ENSMUST00000033054	protein_coding	1,123	0,00223	1,697	0,00001
Adora2b	protein_coding	ENSMUST00000018644	protein_coding	1,462	0,00014	1,940	0,00001
Aen	protein_coding	ENSMUST00000107425	protein_coding	2,309	0,00169	2,128	0,00000
Aen	protein_coding	ENSMUST00000107421	protein_coding	1,817	0,00005	1,672	0,00002
Agap3	protein_coding	ENSMUST00000199856	protein_coding	1,203	0,00644	1,129	0,00002
Akr1b8	protein_coding	ENSMUST00000038406	protein_coding	1,791	0,00006	3,816	0,00000
Aldh1a1	protein_coding	ENSMUST00000087638	protein_coding	1,855	0,00060	2,619	0,00002
Aldh1a7	protein_coding	ENSMUST00000025656	protein_coding	1,669	0,00159	1,460	0,00010
Ankrd1	protein_coding	ENSMUST00000025718	protein_coding	1,177	0,00094	2,236	0,00022
Apobec3	protein_coding	ENSMUST00000175849	retained_intron	1,301	0,00646	1,299	0,00001
Apoc3	protein_coding	ENSMUST00000034586	protein_coding	1,075	0,00117	1,822	0,00004
Arap1	protein_coding	ENSMUST00000133423	protein_coding	3,726	0,00681	3,634	0,00028
Areg	protein_coding	ENSMUST00000031325	protein_coding	3,668	0,00412	5,125	0,00006
Arg2	protein_coding	ENSMUST00000021550	protein_coding	1,659	0,00006	3,684	0,00000
Asf1b	protein_coding	ENSMUST00000005607	protein_coding	1,155	0,00956	1,495	0,00423
Asns	protein_coding	ENSMUST00000031766	protein_coding	1,348	0,00033	1,505	0,00090
Atf3	protein_coding	ENSMUST00000027941	protein_coding	2,755	0,00135	3,164	0,00001
Atg9b	protein_coding	ENSMUST00000059401	protein_coding	2,839	0,00117	4,699	0,00001
Atp13a3	protein_coding	ENSMUST00000229503	protein_coding	5,397	0,00027	4,747	0,00553
Bag2	protein_coding	ENSMUST00000044691	protein_coding	1,256	0,00015	1,042	0,00000

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Bax	protein_coding	ENSMUST00000210392	protein_coding	1,585	0,00010	1,277	0,00002
Bax	protein_coding	ENSMUST00000033093	protein_coding	1,390	0,00008	1,162	0,00000
Bax	protein_coding	ENSMUST00000211365	protein_coding	1,323	0,00294	1,732	0,00016
Bbc3	protein_coding	ENSMUST00000136781	protein_coding	4,131	0,00296	2,021	0,00873
Bbc3	protein_coding	ENSMUST00000002152	protein_coding	2,763	0,00005	2,440	0,00002
Bcl211	protein_coding	ENSMUST00000140436	protein_coding	1,821	0,00292	1,235	0,00023
Bcl3	protein_coding	ENSMUST00000120537	protein_coding	1,512	0,00008	1,272	0,00024
Bra1	protein_coding	ENSMUST00000017290	protein_coding	1,778	0,00029	1,732	0,00045
Bspry	protein_coding	ENSMUST00000030088	protein_coding	1,717	0,00085	1,086	0,00018
Btg2	protein_coding	ENSMUST00000020692	protein_coding	2,625	0,00008	2,938	0,00000
C3	protein_coding	ENSMUST00000024988	protein_coding	2,367	0,00010	3,286	0,00001
Car13	protein_coding	ENSMUST00000029071	protein_coding	1,387	0,00022	1,772	0,00000
Card14	protein_coding	ENSMUST00000151052	processed_transcript	1,330	0,00459	1,224	0,00612
Card14	protein_coding	ENSMUST00000053245	protein_coding	1,247	0,00427	1,695	0,00002
Cast	protein_coding	ENSMUST00000065629	protein_coding	1,400	0,00987	1,112	0,00004
Cbr1	protein_coding	ENSMUST00000039659	protein_coding	1,081	0,00024	1,918	0,00002
Cbr3	protein_coding	ENSMUST00000039620	protein_coding	2,108	0,00010	3,163	0,00000
Ccdc120	protein_coding	ENSMUST00000033490	protein_coding	1,276	0,00017	1,360	0,00003
Ccnf	protein_coding	ENSMUST00000115390	protein_coding	1,970	0,00544	2,898	0,00001
Ccng1	protein_coding	ENSMUST00000020576	protein_coding	1,823	0,00002	1,391	0,00000
Cd14	protein_coding	ENSMUST00000061829	protein_coding	1,594	0,00026	2,583	0,00000
Cd151	protein_coding	ENSMUST00000058746	protein_coding	1,861	0,00155	1,126	0,00018
Cd151	protein_coding	ENSMUST00000210308	retained_intron	1,782	0,00367	1,608	0,00050
Cd68	protein_coding	ENSMUST00000018918	protein_coding	2,836	0,00016	3,265	0,00000
Cdh24	protein_coding	ENSMUST00000067784	protein_coding	1,624	0,00060	1,125	0,00003
Cdkn1a	protein_coding	ENSMUST00000122348	protein_coding	4,640	0,00109	2,575	0,00004
Cdkn1a	protein_coding	ENSMUST00000023829	protein_coding	4,616	0,00003	5,323	0,00000
Cdkn1a	protein_coding	ENSMUST00000119901	protein_coding	4,549	0,00500	5,025	0,00001
Cdt1	protein_coding	ENSMUST00000006760	protein_coding	1,001	0,00127	2,741	0,00230
Cebpb	protein_coding	ENSMUST00000070642	protein_coding	1,508	0,00098	1,763	0,00001
Celf5	protein_coding	ENSMUST00000118763	protein_coding	4,459	0,00934	4,420	0,00228
Cenpt	protein_coding	ENSMUST00000040776	protein_coding	1,030	0,00077	1,636	0,00003
Cep170b	protein_coding	ENSMUST00000222532	retained_intron	1,220	0,00254	1,046	0,00007
Cercam	protein_coding	ENSMUST00000047521	protein_coding	1,436	0,00050	1,954	0,00000
Ces2e	protein_coding	ENSMUST00000109410	protein_coding	2,879	0,00009	2,936	0,00000
Ces2e	protein_coding	ENSMUST00000034355	protein_coding	2,378	0,00009	1,550	0,00031
Cfap157	protein_coding	ENSMUST00000102813	protein_coding	1,808	0,00053	2,270	0,00001
Cgn	protein_coding	ENSMUST00000107272	protein_coding	1,238	0,00075	1,233	0,00051
Chac1	protein_coding	ENSMUST00000028780	protein_coding	1,155	0,00162	1,283	0,00020
Ckap2	protein_coding	ENSMUST00000046916	protein_coding	3,850	0,00016	3,423	0,00002
Ckap2	protein_coding	ENSMUST00000211629	retained_intron	2,963	0,00124	2,537	0,00899
Clefl	protein_coding	ENSMUST00000132305	retained_intron	1,640	0,00008	2,936	0,00000
Cldn23	protein_coding	ENSMUST00000060128	protein_coding	2,178	0,00075	3,989	0,00072
Cldn4	protein_coding	ENSMUST00000051401	protein_coding	1,183	0,00033	2,912	0,00000
Cox6b2	protein_coding	ENSMUST00000182272	retained_intron	1,466	0,00115	2,160	0,00002
Cp	protein_coding	ENSMUST00000150264	retained_intron	1,249	0,00058	1,896	0,00032
Cpne5	protein_coding	ENSMUST00000126993	retained_intron	2,944	0,00039	3,568	0,00080

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Cpt1c	protein_coding	ENSMUST00000211901	retained_intron	1,448	0,00715	1,777	0,00001
Creb3l3	protein_coding	ENSMUST00000117422	protein_coding	2,536	0,00607	1,615	0,00063
Cryab	protein_coding	ENSMUST00000216755	protein_coding	1,182	0,00034	2,050	0,00005
Cstb	protein_coding	ENSMUST00000005185	protein_coding	1,219	0,00009	1,455	0,00000
Ctxn1	protein_coding	ENSMUST00000053252	protein_coding	1,305	0,00054	1,679	0,00004
Cxadr	protein_coding	ENSMUST00000114229	protein_coding	1,469	0,00325	1,139	0,00032
Cxcl1	protein_coding	ENSMUST00000201245	retained_intron	4,526	0,00034	2,837	0,00005
Cxcl1	protein_coding	ENSMUST00000031327	protein_coding	3,754	0,00009	3,452	0,00001
Cxcl10	protein_coding	ENSMUST00000038816	protein_coding	2,691	0,00211	1,407	0,00007
Cyp2s1	protein_coding	ENSMUST00000043314	protein_coding	1,022	0,00009	1,470	0,00000
Cyp4f17	protein_coding	ENSMUST00000165999	protein_coding	1,039	0,00014	1,735	0,00000
Cyr61	protein_coding	ENSMUST00000029846	protein_coding	1,940	0,00009	1,304	0,00069
Deaf4	protein_coding	ENSMUST00000021645	protein_coding	1,014	0,00293	1,317	0,00059
Ddias	protein_coding	ENSMUST00000032877	protein_coding	3,515	0,00006	3,698	0,00000
Ddit4	protein_coding	ENSMUST00000020308	protein_coding	1,166	0,00045	1,053	0,00382
Ddit4l	protein_coding	ENSMUST00000053855	protein_coding	1,266	0,00007	1,491	0,00001
Dhrs9	protein_coding	ENSMUST00000063690	protein_coding	2,852	0,00519	3,200	0,00001
Dnase2a	protein_coding	ENSMUST00000148573	processed_transcript	1,079	0,00039	1,073	0,00013
Dtx4	protein_coding	ENSMUST00000045521	protein_coding	1,860	0,00013	1,233	0,00003
Dusp5	protein_coding	ENSMUST00000038287	protein_coding	1,733	0,00045	3,287	0,00000
Dusp8	protein_coding	ENSMUST00000039926	protein_coding	1,811	0,00035	1,764	0,00000
Dyrk3	protein_coding	ENSMUST00000016670	protein_coding	1,735	0,00047	1,818	0,00000
Eda2r	protein_coding	ENSMUST00000126283	retained_intron	4,199	0,00050	4,149	0,00000
Eda2r	protein_coding	ENSMUST00000037353	protein_coding	3,863	0,00005	4,372	0,00000
Efh2	protein_coding	ENSMUST00000036854	protein_coding	1,095	0,00011	1,599	0,00000
Egr1	protein_coding	ENSMUST00000064795	protein_coding	2,915	0,00215	1,618	0,00593
Eid3	protein_coding	ENSMUST00000211623	protein_coding	2,253	0,00329	1,636	0,00000
Enc1	protein_coding	ENSMUST00000041623	protein_coding	1,606	0,00022	2,312	0,00000
Epha2	protein_coding	ENSMUST00000006614	protein_coding	1,522	0,00055	2,528	0,00000
Ephx1	protein_coding	ENSMUST00000036928	protein_coding	1,299	0,00008	1,839	0,00000
Eppk1	protein_coding	ENSMUST00000226781	protein_coding	1,044	0,00119	1,650	0,00001
Eps8l3	protein_coding	ENSMUST00000037375	protein_coding	2,155	0,00079	2,130	0,00133
Erccl	protein_coding	ENSMUST00000003645	protein_coding	1,502	0,00118	2,536	0,00000
Ermard	protein_coding	ENSMUST00000226213	protein_coding	1,208	0,00111	1,117	0,00225
Fam57a	protein_coding	ENSMUST00000094014	protein_coding	1,176	0,00149	1,286	0,00007
Fam83c	protein_coding	ENSMUST00000029143	protein_coding	1,970	0,00285	1,376	0,00002
Fat1	protein_coding	ENSMUST00000191428	protein_coding	1,841	0,00149	1,093	0,00000
Fblim1	protein_coding	ENSMUST00000105784	protein_coding	1,192	0,00834	1,998	0,00673
Fbp2	protein_coding	ENSMUST00000021907	protein_coding	1,151	0,00084	1,014	0,00013
Fbxw7	protein_coding	ENSMUST00000029727	protein_coding	3,827	0,00795	1,958	0,00003
Fbxw9	protein_coding	ENSMUST00000141183	retained_intron	1,750	0,00060	2,038	0,00436
Fbxw9	protein_coding	ENSMUST00000139721	retained_intron	1,343	0,00691	1,479	0,00867
Fbxw9	protein_coding	ENSMUST00000095220	protein_coding	1,172	0,00066	1,031	0,00001
Fga	protein_coding	ENSMUST00000029630	protein_coding	1,748	0,00006	3,651	0,00000
Fga	protein_coding	ENSMUST00000166581	protein_coding	1,647	0,00751	3,002	0,00026
Fgd3	protein_coding	ENSMUST00000110087	protein_coding	1,196	0,00031	1,274	0,00004
Fgd3	protein_coding	ENSMUST00000048716	protein_coding	1,177	0,00140	1,174	0,00027

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Fjx1	protein_coding	ENSMUST00000099678	protein_coding	1,466	0,00056	1,278	0,00009
Fosb	protein_coding	ENSMUST00000003640	protein_coding	5,163	0,00142	5,047	0,00294
Fosl1	protein_coding	ENSMUST00000025850	protein_coding	3,245	0,00005	3,978	0,00000
Foxj1	protein_coding	ENSMUST00000036215	protein_coding	3,742	0,00004	2,520	0,00005
Fyb2	protein_coding	ENSMUST00000106803	protein_coding	5,165	0,00792	5,494	0,00271
Fyb2	protein_coding	ENSMUST00000106804	protein_coding	1,502	0,00047	1,035	0,00250
G6pdx	protein_coding	ENSMUST00000004327	protein_coding	1,140	0,00005	1,090	0,00001
Gadd45b	protein_coding	ENSMUST00000015456	protein_coding	1,060	0,00029	2,167	0,00003
Galr3	protein_coding	ENSMUST00000058004	protein_coding	1,695	0,00508	1,930	0,00008
Gar1	protein_coding	ENSMUST00000029643	protein_coding	1,086	0,00320	1,008	0,00514
Gas211	protein_coding	ENSMUST00000122886	processed_transcript	1,543	0,00591	1,577	0,00024
Gc	protein_coding	ENSMUST00000200534	protein_coding	3,326	0,00141	1,838	0,00025
Gc	protein_coding	ENSMUST00000049209	protein_coding	2,928	0,00003	3,097	0,00000
Gcnt2	protein_coding	ENSMUST00000067778	protein_coding	1,319	0,00736	2,825	0,00177
Gdf15	protein_coding	ENSMUST00000110103	protein_coding	3,308	0,00007	3,509	0,00000
Gm1070	lncRNA	ENSMUST00000114185	lncRNA	2,667	0,00482	2,528	0,00493
Gm13836	pseudogene	ENSMUST00000121159	processed_pseudogene	3,727	0,00662	3,245	0,00077
Gm15264	pseudogene	ENSMUST00000118050	unprocessed_pseudogene	1,127	0,00323	1,521	0,00404
Gm16098	lncRNA	ENSMUST00000141926	antisense	2,206	0,00941	3,719	0,00023
Gm17501	lncRNA	ENSMUST00000181247	lncRNA	2,246	0,00105	2,057	0,00002
Gm20186	lncRNA	ENSMUST00000201775	lncRNA	1,731	0,00860	2,131	0,00001
Gm26525	lncRNA	ENSMUST00000180666	lncRNA	1,132	0,00472	1,046	0,00125
Gm26669	lncRNA	ENSMUST00000181323	lncRNA	1,452	0,00311	1,204	0,00017
Gm26674	lncRNA	ENSMUST00000180770	antisense	1,574	0,00111	1,007	0,00033
Gm26917	lncRNA	ENSMUST00000192833	lncRNA	2,308	0,00025	1,389	0,00000
Gm32200	lncRNA	ENSMUST00000192778	lncRNA	2,532	0,00894	1,421	0,00010
Gm3776	protein_coding	ENSMUST00000217203	protein_coding	7,088	0,00010	6,118	0,00001
Gm37795	TEC	ENSMUST00000194216	TEC	2,225	0,00006	1,627	0,00000
Gm39460	lncRNA	ENSMUST00000215728	lncRNA	4,446	0,00281	2,690	0,00011
Gm42793	TEC	ENSMUST00000199139	TEC	1,831	0,00458	2,823	0,00000
Gm44911	pseudogene	ENSMUST00000208436	processed_pseudogene	2,265	0,00113	3,973	0,00045
Gm47430	lncRNA	ENSMUST00000214586	lncRNA	2,810	0,00175	2,118	0,00076
Gm49141	pseudogene	ENSMUST00000223731	processed_transcript	1,637	0,00109	4,664	0,00001
Gm49146	pseudogene	ENSMUST00000225166	processed_transcript	2,094	0,00033	3,767	0,00002
Gm49146	pseudogene	ENSMUST00000223745	processed_transcript	1,772	0,00029	3,423	0,00001
Gm5799	protein_coding	ENSMUST00000169023	protein_coding	2,200	0,00040	4,867	0,00005
Gm8126	protein_coding	ENSMUST00000169587	protein_coding	2,185	0,00006	3,880	0,00001
Gm9530	lncRNA	ENSMUST00000181160	lncRNA	2,574	0,00085	1,198	0,00015
Gm9732	protein_coding	ENSMUST00000228840	processed_transcript	1,727	0,00209	3,568	0,00009
Gm9949	lncRNA	ENSMUST00000067743	protein_coding	1,683	0,00175	1,570	0,00005
Gsta1	protein_coding	ENSMUST00000098537	protein_coding	3,241	0,00008	7,038	0,00002
Gtse1	protein_coding	ENSMUST00000231074	protein_coding	6,971	0,00006	3,924	0,00800
Gtse1	protein_coding	ENSMUST00000170629	protein_coding	2,933	0,00057	5,773	0,00177
Havcr1	protein_coding	ENSMUST00000109223	protein_coding	1,525	0,00006	6,322	0,00001
Havcr1	protein_coding	ENSMUST00000047568	protein_coding	1,511	0,00006	8,000	0,00000
Hcn3	protein_coding	ENSMUST00000029686	protein_coding	1,453	0,00076	1,379	0,00003
Hmga1b	protein_coding	ENSMUST00000105046	protein_coding	1,007	0,00015	3,996	0,00063

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Hmox1	protein_coding	ENSMUST00000005548	protein_coding	1,995	0,00005	2,007	0,00000
Homer3	protein_coding	ENSMUST00000155711	retained_intron	2,021	0,00374	2,608	0,00002
Hpgds	protein_coding	ENSMUST00000031982	protein_coding	1,607	0,00034	1,627	0,00017
Hspb1	protein_coding	ENSMUST00000005077	protein_coding	1,561	0,00010	2,117	0,00004
Icam1	protein_coding	ENSMUST00000086399	protein_coding	1,466	0,00008	1,422	0,00001
Ier2	protein_coding	ENSMUST00000060427	protein_coding	1,229	0,00156	1,421	0,00094
Ier5	protein_coding	ENSMUST00000055322	protein_coding	1,694	0,00013	2,003	0,00000
Iifi2712b	protein_coding	ENSMUST00000044687	protein_coding	3,197	0,00808	3,389	0,00007
Iifrd1	protein_coding	ENSMUST00000171553	retained_intron	1,792	0,00030	2,933	0,00001
Iifrd1	protein_coding	ENSMUST00000165027	protein_coding	1,518	0,00008	2,240	0,00000
Iifrd1	protein_coding	ENSMUST00000164047	protein_coding	1,332	0,00020	2,162	0,00007
Il1m	protein_coding	ENSMUST00000114487	protein_coding	4,013	0,00399	6,339	0,00000
Inava	protein_coding	ENSMUST00000144464	protein_coding	1,839	0,00027	1,622	0,00000
Inava	protein_coding	ENSMUST00000120339	protein_coding	1,838	0,00005	1,623	0,00000
Inka2	protein_coding	ENSMUST00000066610	protein_coding	4,039	0,00008	2,668	0,00000
Inka2	protein_coding	ENSMUST00000098273	protein_coding	3,121	0,00153	1,609	0,00001
Jpt1	protein_coding	ENSMUST00000021083	protein_coding	1,018	0,00013	1,367	0,00000
Junb	protein_coding	ENSMUST00000064922	protein_coding	2,095	0,00012	1,850	0,00024
Kctd11	protein_coding	ENSMUST00000050555	protein_coding	1,111	0,00017	1,260	0,00001
Kdelr3	protein_coding	ENSMUST00000010974	protein_coding	1,365	0,00415	1,543	0,00002
Klfb	protein_coding	ENSMUST00000000080	protein_coding	1,595	0,00032	1,714	0,00000
Klrg2	protein_coding	ENSMUST00000096030	protein_coding	2,340	0,00283	5,542	0,00016
Krt20	protein_coding	ENSMUST00000017743	protein_coding	5,931	0,00006	8,128	0,00027
Lcn2	protein_coding	ENSMUST00000050785	protein_coding	1,904	0,00105	6,016	0,00000
Lgals1	protein_coding	ENSMUST00000089377	protein_coding	1,335	0,00074	1,341	0,00000
Lif	protein_coding	ENSMUST00000040750	protein_coding	5,395	0,00432	5,245	0,00289
Lif	protein_coding	ENSMUST00000066283	protein_coding	2,527	0,00149	4,132	0,00850
Lmna	protein_coding	ENSMUST00000120377	protein_coding	1,196	0,00033	1,250	0,00000
Ltb4r1	protein_coding	ENSMUST00000057569	protein_coding	1,632	0,00264	2,132	0,00011
Ltb4r2	protein_coding	ENSMUST00000228302	retained_intron	1,780	0,00094	1,266	0,00024
Mab2113	protein_coding	ENSMUST00000061831	protein_coding	2,960	0,00071	1,543	0,00012
Mab2113	protein_coding	ENSMUST00000118411	protein_coding	2,222	0,00010	2,315	0,00001
Maff	protein_coding	ENSMUST00000096350	protein_coding	3,195	0,00394	3,513	0,00004
Map3k6	protein_coding	ENSMUST00000030677	protein_coding	1,259	0,00525	1,979	0,00001
Mast1	protein_coding	ENSMUST00000153000	retained_intron	1,055	0,00214	1,292	0,00009
Mdm2	protein_coding	ENSMUST00000155285	nonsense_mediated_decay	1,847	0,00009	6,217	0,00425
Mdm2	protein_coding	ENSMUST00000147823	processed_transcript	1,592	0,00706	1,403	0,00011
Mdm2	protein_coding	ENSMUST00000131627	retained_intron	1,583	0,00038	2,275	0,00000
Mdm2	protein_coding	ENSMUST00000020408	protein_coding	1,374	0,00074	1,431	0,00001
Mdm2	protein_coding	ENSMUST00000132277	retained_intron	1,280	0,00014	1,711	0,00000
Mesp2	protein_coding	ENSMUST00000107394	protein_coding	3,180	0,00282	1,899	0,00011
Mettl21a	protein_coding	ENSMUST00000162806	processed_transcript	4,152	0,00276	2,016	0,00500
Mettl21a	protein_coding	ENSMUST00000114079	protein_coding	1,430	0,00042	1,059	0,00003
Mgst1	protein_coding	ENSMUST00000120302	protein_coding	1,612	0,00007	1,188	0,00022
Mgst1	protein_coding	ENSMUST00000008684	protein_coding	1,521	0,00011	1,167	0,00037
Mid1ip1	protein_coding	ENSMUST00000008179	protein_coding	1,193	0,00016	1,162	0,00008
Mirt1	lincRNA	ENSMUST00000180489	lincRNA	2,054	0,00123	3,064	0,00000

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Mirt1	lncRNA	ENSMUST00000180667	lincRNA	1,544	0,00369	2,383	0,00015
Morf412	protein_coding	ENSMUST00000169418	protein_coding	5,503	0,00471	5,109	0,00065
Mroh3	protein_coding	ENSMUST00000146112	processed_transcript	3,616	0,00663	3,211	0,00219
Mroh3	protein_coding	ENSMUST00000168561	nonsense_mediated_decay	2,988	0,00757	4,788	0,00047
Mt2	protein_coding	ENSMUST00000034214	protein_coding	1,115	0,00462	3,699	0,00000
Mthfd2	protein_coding	ENSMUST00000005810	protein_coding	1,555	0,00010	1,352	0,00002
Mybl1	protein_coding	ENSMUST00000088658	protein_coding	2,532	0,00011	3,282	0,00000
Myc	protein_coding	ENSMUST00000160009	protein_coding	2,085	0,00072	1,891	0,00789
Myc	protein_coding	ENSMUST00000022971	protein_coding	2,023	0,00149	2,221	0,00107
Myc	protein_coding	ENSMUST00000161976	protein_coding	2,023	0,00149	2,221	0,00107
Myl6	protein_coding	ENSMUST00000220307	nonsense_mediated_decay	1,312	0,00606	1,365	0,00005
Myof	protein_coding	ENSMUST00000226068	protein_coding	1,533	0,00079	2,159	0,00011
Nabp1	protein_coding	ENSMUST00000189542	nonsense_mediated_decay	1,651	0,00170	1,367	0,00015
Nabp1	protein_coding	ENSMUST00000186003	nonsense_mediated_decay	1,541	0,00010	1,187	0,00005
Nanos1	protein_coding	ENSMUST00000088237	protein_coding	2,584	0,00079	1,431	0,00010
Nap111	protein_coding	ENSMUST00000065917	protein_coding	1,229	0,00149	1,574	0,00000
Nek6	protein_coding	ENSMUST00000112895	protein_coding	1,305	0,00956	1,680	0,00001
Nek6	protein_coding	ENSMUST00000112902	protein_coding	1,099	0,00010	1,888	0,00000
Nfkb2	protein_coding	ENSMUST00000073116	protein_coding	1,620	0,00014	1,678	0,00000
Nfkb2	protein_coding	ENSMUST00000111881	protein_coding	1,003	0,00928	1,527	0,00036
Nfkbiz	protein_coding	ENSMUST00000114458	protein_coding	1,433	0,00016	1,837	0,00000
Nqo1	protein_coding	ENSMUST00000003947	protein_coding	1,619	0,00002	1,455	0,00001
Nrg1	protein_coding	ENSMUST00000208617	protein_coding	1,876	0,00359	2,191	0,00001
Nrg1	protein_coding	ENSMUST00000207584	retained_intron	1,860	0,00153	2,026	0,00012
Nrg1	protein_coding	ENSMUST00000207470	protein_coding	1,417	0,00103	1,244	0,00001
Pdgfa	protein_coding	ENSMUST00000144934	processed_transcript	4,692	0,00376	1,377	0,00354
Pdgfa	protein_coding	ENSMUST00000046901	protein_coding	1,142	0,00014	1,136	0,00000
Pdgfa	protein_coding	ENSMUST00000076095	protein_coding	1,042	0,00315	1,090	0,00099
Pdk4	protein_coding	ENSMUST00000019721	protein_coding	2,236	0,00091	3,245	0,00009
Pdlim7	protein_coding	ENSMUST00000069968	protein_coding	4,538	0,00745	3,516	0,00656
Pgd	protein_coding	ENSMUST00000084124	protein_coding	1,202	0,00005	1,295	0,00000
Pgd	protein_coding	ENSMUST00000124409	retained_intron	1,072	0,00033	1,472	0,00001
Pgd	protein_coding	ENSMUST00000156120	retained_intron	1,049	0,00924	1,814	0,00096
Phlda1	protein_coding	ENSMUST00000164773	protein_coding	1,157	0,00162	1,884	0,00000
Phlda3	protein_coding	ENSMUST00000038945	protein_coding	3,746	0,00002	3,125	0,00000
Plaur	protein_coding	ENSMUST00000002284	protein_coding	1,342	0,00084	2,863	0,00000
Plec	protein_coding	ENSMUST00000072692	protein_coding	6,536	0,00167	1,612	0,00001
Plin2	protein_coding	ENSMUST00000134437	processed_transcript	3,670	0,00822	3,115	0,00000
Plin2	protein_coding	ENSMUST00000000466	protein_coding	3,273	0,00002	2,686	0,00000
Plin2	protein_coding	ENSMUST00000134094	processed_transcript	3,181	0,00127	2,831	0,00000
Plin2	protein_coding	ENSMUST00000154999	processed_transcript	2,894	0,00132	2,210	0,00001
Plk2	protein_coding	ENSMUST00000022212	protein_coding	2,460	0,00009	2,382	0,00000
Plk3	protein_coding	ENSMUST00000076859	protein_coding	1,813	0,00087	1,218	0,00004
Pmaip1	protein_coding	ENSMUST00000025399	protein_coding	1,642	0,00106	1,568	0,00000
Pmm1	protein_coding	ENSMUST00000129406	retained_intron	1,331	0,00181	1,952	0,00030
Pmm1	protein_coding	ENSMUST00000023112	protein_coding	1,165	0,00009	1,389	0,00000
Pnrc1	protein_coding	ENSMUST00000155616	processed_transcript	2,268	0,00078	1,583	0,00003

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Polk	protein_coding	ENSMUST00000022172	protein_coding	1,150	0,00014	1,147	0,00000
Ppm1j	protein_coding	ENSMUST00000002298	protein_coding	1,496	0,00041	2,225	0,00000
Ppp1r13l	protein_coding	ENSMUST00000047621	protein_coding	1,294	0,00013	2,581	0,00000
Praf2	protein_coding	ENSMUST00000033489	protein_coding	1,053	0,00018	1,234	0,00007
Prrg4	protein_coding	ENSMUST00000028593	protein_coding	2,519	0,00030	2,523	0,00000
Prss56	protein_coding	ENSMUST00000044533	protein_coding	3,611	0,00521	5,290	0,00000
Psd	protein_coding	ENSMUST00000226062	retained_intron	1,671	0,00085	1,485	0,00003
Psd	protein_coding	ENSMUST00000225748	processed_transcript	1,113	0,00314	1,654	0,00000
Psrc1	protein_coding	ENSMUST00000128089	protein_coding	4,143	0,00286	4,217	0,00007
Psrc1	protein_coding	ENSMUST00000128962	processed_transcript	4,004	0,00313	3,820	0,00058
Psrc1	protein_coding	ENSMUST00000090561	protein_coding	3,865	0,00011	5,362	0,00047
Psrc1	protein_coding	ENSMUST00000196430	retained_intron	3,645	0,00127	2,206	0,00000
Ptp4a1	protein_coding	ENSMUST00000153681	retained_intron	1,933	0,00756	1,398	0,00043
Ptp4a1	protein_coding	ENSMUST00000076587	protein_coding	1,185	0,00007	1,547	0,00000
Ptrh1	protein_coding	ENSMUST00000138214	retained_intron	2,410	0,00790	2,151	0,00225
Ptrh1	protein_coding	ENSMUST00000066352	protein_coding	2,113	0,00028	2,114	0,00001
Pvr	protein_coding	ENSMUST00000043517	protein_coding	1,377	0,00026	1,656	0,00000
Pvr	protein_coding	ENSMUST00000160572	retained_intron	1,290	0,00617	1,870	0,00001
Rab44	protein_coding	ENSMUST00000233717	protein_coding	2,498	0,00804	2,419	0,00358
Ran	protein_coding	ENSMUST00000111343	protein_coding	1,314	0,00498	1,407	0,00284
Rap2b	protein_coding	ENSMUST00000049064	protein_coding	2,357	0,00009	3,036	0,00000
Rcan1	protein_coding	ENSMUST00000023672	protein_coding	1,809	0,00018	1,047	0,00010
Rcn1	protein_coding	ENSMUST00000006128	protein_coding	1,013	0,00043	1,153	0,00004
Rem2	protein_coding	ENSMUST00000164766	protein_coding	1,916	0,00148	2,150	0,00095
Rhbd2	protein_coding	ENSMUST00000103029	protein_coding	1,617	0,00006	1,714	0,00000
Rhbd2	protein_coding	ENSMUST00000126819	retained_intron	1,404	0,00096	1,709	0,00000
Rin1	protein_coding	ENSMUST00000025818	protein_coding	1,530	0,00888	2,592	0,00000
Rnd1	protein_coding	ENSMUST00000003451	protein_coding	1,386	0,00157	2,572	0,00000
Rnf223	protein_coding	ENSMUST00000209248	protein_coding	1,732	0,00475	3,275	0,00049
Rpl10a	protein_coding	ENSMUST00000129935	nonsense_mediated_decay	1,180	0,00639	1,004	0,00076
Rpl12	protein_coding	ENSMUST00000102811	nonsense_mediated_decay	1,315	0,00443	1,850	0,00023
Rpl12	protein_coding	ENSMUST00000154340	retained_intron	1,036	0,00069	1,263	0,00003
Rprm	protein_coding	ENSMUST00000100089	protein_coding	1,849	0,00110	3,191	0,00000
Rsph4a	protein_coding	ENSMUST00000169670	protein_coding	2,915	0,00011	1,590	0,00006
Rusc2	protein_coding	ENSMUST00000135216	protein_coding	3,487	0,00700	1,771	0,00004
Rusc2	protein_coding	ENSMUST00000035645	protein_coding	1,089	0,00574	1,948	0,00013
S100a10	protein_coding	ENSMUST00000148400	retained_intron	1,646	0,00880	1,879	0,00003
S100a10	protein_coding	ENSMUST00000045756	protein_coding	1,583	0,00005	2,248	0,00000
Saa3	protein_coding	ENSMUST00000006956	protein_coding	3,630	0,00550	4,189	0,00011
Saa4	protein_coding	ENSMUST00000006952	protein_coding	4,139	0,00006	1,352	0,00004
Samd5	protein_coding	ENSMUST00000100070	protein_coding	1,124	0,00231	1,542	0,00001
Sbno2	protein_coding	ENSMUST00000219258	nonsense_mediated_decay	1,073	0,00293	1,289	0,00001
Scmh1	protein_coding	ENSMUST00000136801	processed_transcript	1,187	0,00014	1,484	0,00011
Sdad1	protein_coding	ENSMUST00000201143	protein_coding	5,591	0,00023	1,044	0,00170
Sec14l2	protein_coding	ENSMUST00000003681	protein_coding	1,092	0,00026	1,228	0,00002
Sec14l5	protein_coding	ENSMUST00000230616	protein_coding	3,064	0,00055	2,718	0,00001
Serpina10	protein_coding	ENSMUST00000121625	protein_coding	1,752	0,00453	5,463	0,00070

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(continuação)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Serpina7	protein_coding	ENSMUST00000033626	protein_coding	5.126	0.00807	3.766	0.00067
Sesn2	protein_coding	ENSMUST00000030724	protein_coding	1.053	0.00009	1.027	0.00000
Sfn	protein_coding	ENSMUST00000057311	protein_coding	2.554	0.00007	2.897	0.00000
Shisa4	protein_coding	ENSMUST00000041240	protein_coding	1.201	0.00315	1.846	0.00000
Siva1	protein_coding	ENSMUST00000021728	protein_coding	1.304	0.00011	1.524	0.00000
Slc16a6	protein_coding	ENSMUST00000070872	protein_coding	1.304	0.00084	1.241	0.00000
Slc19a2	protein_coding	ENSMUST00000044021	protein_coding	1.612	0.00025	1.551	0.00000
Slc20a1	protein_coding	ENSMUST000000140907	retained_intron	2.018	0.00848	1.032	0.00174
Slc20a1	protein_coding	ENSMUST000000144025	retained_intron	1.792	0.00263	1.436	0.00142
Slc20a1	protein_coding	ENSMUST00000028880	protein_coding	1.540	0.00040	1.293	0.00000
Slc20a1	protein_coding	ENSMUST000000141285	processed_transcript	1.507	0.00134	1.054	0.00067
Slc35e4	protein_coding	ENSMUST00000051207	protein_coding	1.212	0.00087	1.343	0.00000
Slc4a11	protein_coding	ENSMUST000000144945	retained_intron	1.255	0.00056	2.035	0.00157
Slc4a11	protein_coding	ENSMUST000000134647	retained_intron	1.180	0.00044	2.165	0.00006
Slc4a11	protein_coding	ENSMUST00000099362	protein_coding	1.114	0.00017	2.076	0.00013
Smox	protein_coding	ENSMUST00000028806	protein_coding	1.574	0.00083	2.346	0.00002
Smpd13b	protein_coding	ENSMUST00000030709	protein_coding	3.472	0.00006	3.527	0.00000
Snai3	protein_coding	ENSMUST00000006762	protein_coding	2.189	0.00132	2.686	0.00000
Snhg15	lncRNA	ENSMUST000000147762	processed_transcript	1.789	0.00118	1.609	0.00001
Snhg15	lncRNA	ENSMUST000000134527	processed_transcript	1.490	0.00136	1.527	0.00020
Snhg15	lncRNA	ENSMUST000000129570	processed_transcript	1.462	0.00244	1.347	0.00014
Socs3	protein_coding	ENSMUST00000054002	protein_coding	1.593	0.00148	2.846	0.00000
Sprr1a	protein_coding	ENSMUST00000054599	protein_coding	1.548	0.00034	4.758	0.00012
Srxn1	protein_coding	ENSMUST00000041500	protein_coding	1.944	0.00002	2.713	0.00001
Stra6l	protein_coding	ENSMUST000000128947	processed_transcript	2.869	0.00123	1.859	0.00491
Tap1	protein_coding	ENSMUST000000170086	protein_coding	1.552	0.00268	1.447	0.00000
Tes	protein_coding	ENSMUST000000115467	protein_coding	1.127	0.00008	1.287	0.00000
Tgif1	protein_coding	ENSMUST00000059775	protein_coding	1.314	0.00030	2.227	0.00000
Thyn1	protein_coding	ENSMUST000000213770	protein_coding	1.520	0.00070	1.785	0.00001
Thyn1	protein_coding	ENSMUST00000039161	protein_coding	1.064	0.00014	1.391	0.00000
Tinagl1	protein_coding	ENSMUST000000175822	processed_transcript	1.145	0.00403	1.100	0.00009
Tmbim1	protein_coding	ENSMUST000000016309	protein_coding	1.176	0.00006	1.496	0.00000
Tmcc3	protein_coding	ENSMUST000000065060	protein_coding	1.017	0.00021	1.328	0.00000
Tmem43	protein_coding	ENSMUST000000032183	protein_coding	1.392	0.00005	1.712	0.00000
Tnfrsf10b	protein_coding	ENSMUST000000222663	protein_coding	2.417	0.00006	2.398	0.00000
Tnfrsf12a	protein_coding	ENSMUST00000024698	protein_coding	2.770	0.00005	2.867	0.00000
Tnfrsf12a	protein_coding	ENSMUST000000233040	retained_intron	2.700	0.00421	3.296	0.00476
Tnfrsf12a	protein_coding	ENSMUST000000233754	retained_intron	2.196	0.00027	2.707	0.00027
Tnfrsf1b	protein_coding	ENSMUST00000030336	protein_coding	1.019	0.00419	1.441	0.00000
Tnfrsf23	protein_coding	ENSMUST000000035742	nonsense_mediated_decay	1.067	0.00809	3.406	0.00000
Tnip1	protein_coding	ENSMUST000000102730	protein_coding	1.189	0.00546	1.652	0.00033
Tor3a	protein_coding	ENSMUST000000122242	nonsense_mediated_decay	4.332	0.00662	1.427	0.00022
Tor3a	protein_coding	ENSMUST000000188964	protein_coding	4.332	0.00662	1.427	0.00022
Trib3	protein_coding	ENSMUST000000040312	protein_coding	1.257	0.00124	1.973	0.00003
Trim6	protein_coding	ENSMUST000000098180	protein_coding	2.300	0.00013	2.692	0.00000
Trp53	protein_coding	ENSMUST00000005371	protein_coding	1.522	0.00155	1.005	0.00001
Trp53cor1	lncRNA	ENSMUST000000133221	lincRNA	3.066	0.00867	2.950	0.00001

**APÊNDICE B – Transcritos superexpressos em comum ao Estudo 1 e Estudo 2
(final)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Trp53inp1	protein_coding	ENSMUST00000029865	protein_coding	2,598	0,00070	2,121	0,00000
Trp53inp1	protein_coding	ENSMUST00000132693	processed_transcript	2,215	0,00118	1,910	0,00010
Trp53inp1	protein_coding	ENSMUST00000132010	processed_transcript	2,189	0,00031	2,206	0,00001
Trpv6	protein_coding	ENSMUST00000031902	protein_coding	1,440	0,00030	1,546	0,00001
Tsc22d1	protein_coding	ENSMUST00000134109	protein_coding	1,314	0,00116	1,001	0,00012
Tsc22d1	protein_coding	ENSMUST00000022587	protein_coding	1,145	0,00030	1,517	0,00000
Ttc39a	protein_coding	ENSMUST00000064129	protein_coding	1,684	0,00849	1,184	0,00007
Tubb6	protein_coding	ENSMUST00000001513	protein_coding	1,928	0,00006	2,072	0,00000
Txnrd1	protein_coding	ENSMUST00000219962	protein_coding	2,670	0,00351	1,389	0,00000
Txnrd1	protein_coding	ENSMUST00000219368	protein_coding	1,892	0,00628	1,846	0,00007
Txnrd1	protein_coding	ENSMUST00000219911	retained_intron	1,700	0,00046	1,485	0,00000
Txnrd1	protein_coding	ENSMUST00000020484	protein_coding	1,584	0,00004	1,289	0,00000
Txnrd1	protein_coding	ENSMUST00000219442	protein_coding	1,488	0,00156	1,049	0,00003
Ucp2	protein_coding	ENSMUST00000149808	processed_transcript	1,055	0,00733	2,262	0,00004
Ucp2	protein_coding	ENSMUST00000126381	retained_intron	1,038	0,00185	2,153	0,00005
Ugdh	protein_coding	ENSMUST00000125375	retained_intron	1,124	0,00015	1,168	0,00001
Ugt1a10	protein_coding	ENSMUST00000113142	protein_coding	1,423	0,00323	1,521	0,00361
Vat1	protein_coding	ENSMUST00000040430	protein_coding	1,601	0,00006	1,514	0,00001
Vmp1	protein_coding	ENSMUST00000145846	processed_transcript	2,059	0,00006	2,288	0,00001
Wnt9a	protein_coding	ENSMUST00000108783	protein_coding	1,741	0,00275	1,773	0,00000
Xdh	protein_coding	ENSMUST00000024866	protein_coding	5,799	0,00168	1,698	0,00000
Ypel2	protein_coding	ENSMUST00000018571	protein_coding	1,544	0,00011	1,239	0,00001
Zbtb42	protein_coding	ENSMUST00000169593	protein_coding	1,188	0,00133	1,467	0,00001
Zfp365	protein_coding	ENSMUST00000064656	protein_coding	2,912	0,00027	4,526	0,00001
Zfp3611	protein_coding	ENSMUST00000021552	protein_coding	1,128	0,00033	1,581	0,00000
Zfp385a	protein_coding	ENSMUST00000229551	protein_coding	2,107	0,00017	2,080	0,00066
Zfp750	protein_coding	ENSMUST00000092298	protein_coding	1,496	0,00018	1,308	0,00000
Zmat3	protein_coding	ENSMUST00000029199	protein_coding	1,690	0,00074	1,774	0,00000

APÊNDICE C – TRANSCRITOS SUBEXPRESSOS EM COMUM AO ESTUDO 1 E

ESTUDO 2

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
0610005C13Rik	lncRNA	ENSMUST00000210866	antisense	-1,046	0,00146	-2,046	0,00010
0610005C13Rik	lncRNA	ENSMUST00000210383	antisense	-1,254	0,00065	-2,412	0,00000
0610005C13Rik	lncRNA	ENSMUST00000210965	antisense	-1,264	0,00615	-1,205	0,00070
0610005C13Rik	lncRNA	ENSMUST00000210569	antisense	-1,670	0,00037	-3,115	0,00004
0610006L08Rik	lncRNA	ENSMUST00000206672	lincRNA	-1,826	0,00044	-2,719	0,00001
0610006L08Rik	lncRNA	ENSMUST00000205326	lincRNA	-1,833	0,00022	-2,727	0,00001
1810059H22Rik	lncRNA	ENSMUST00000204570	lincRNA	-1,493	0,00013	-1,004	0,00011
2210408I21Rik	protein_coding	ENSMUST00000168779	protein_coding	-2,725	0,00062	-1,448	0,00001
4930434B07Rik	lncRNA	ENSMUST00000190254	lincRNA	-1,514	0,00834	-3,266	0,00078
4930452B06Rik	protein_coding	ENSMUST00000102996	protein_coding	-2,775	0,00279	-1,537	0,00095
4930533I22Rik	lncRNA	ENSMUST00000176290	antisense	-2,212	0,00029	-2,234	0,00002
5730403I07Rik	lncRNA	ENSMUST00000183878	lincRNA	-2,245	0,00272	-1,488	0,00590
6330410L21Rik	lncRNA	ENSMUST00000199846	antisense	-1,117	0,00079	-1,193	0,00001
A4gnt	protein_coding	ENSMUST00000042553	protein_coding	-1,144	0,00017	-2,714	0,00060
A630076J17Rik	protein_coding	ENSMUST00000179399	protein_coding	-1,355	0,00153	-1,917	0,00149
Aacs	protein_coding	ENSMUST00000031445	protein_coding	-1,207	0,00126	-1,635	0,00006
Aass	protein_coding	ENSMUST00000031707	protein_coding	-1,132	0,00043	-1,323	0,00000
Abat	protein_coding	ENSMUST00000065987	protein_coding	-1,124	0,00006	-1,881	0,00000
Abca13	protein_coding	ENSMUST00000042740	protein_coding	-1,737	0,00014	-3,489	0,00000
Abcc6	protein_coding	ENSMUST00000002850	protein_coding	-1,389	0,00011	-2,970	0,00000
AC107792.3	TEC	ENSMUST00000233431	TEC	-1,155	0,00554	-1,228	0,00014
AC114585.1	lncRNA	ENSMUST00000228080	lincRNA	-1,991	0,00229	-2,070	0,00032
Acad10	protein_coding	ENSMUST00000111770	protein_coding	-1,251	0,00006	-1,477	0,00000
Acad9	protein_coding	ENSMUST00000195896	retained_intron	-1,609	0,00265	-1,075	0,00272
Acat1	protein_coding	ENSMUST00000034547	protein_coding	-1,229	0,00013	-2,023	0,00000
Acat1	protein_coding	ENSMUST00000169317	processed_transcript	-1,245	0,00013	-1,368	0,00024
Ace	protein_coding	ENSMUST00000001963	protein_coding	-1,192	0,00005	-1,933	0,00000
Acmsd	protein_coding	ENSMUST00000038006	protein_coding	-1,232	0,00190	-3,605	0,00006
Acot11	protein_coding	ENSMUST00000102762	protein_coding	-1,176	0,00206	-1,387	0,00038
Acox2	protein_coding	ENSMUST00000171567	processed_transcript	-1,476	0,00060	-1,226	0,00006
Acox3	protein_coding	ENSMUST00000156125	protein_coding	-1,048	0,00640	-1,807	0,00009
Acox3	protein_coding	ENSMUST00000068947	protein_coding	-1,218	0,00013	-1,500	0,00000
Acox3	protein_coding	ENSMUST00000114237	protein_coding	-1,412	0,00066	-1,843	0,00005
Acsm1	protein_coding	ENSMUST00000047929	protein_coding	-1,143	0,00006	-1,609	0,00001
Acsm2	protein_coding	ENSMUST00000167935	protein_coding	-1,145	0,00012	-2,851	0,00000
Acsm2	protein_coding	ENSMUST00000098084	protein_coding	-1,191	0,00090	-3,167	0,00000
Acsm2	protein_coding	ENSMUST00000129766	protein_coding	-1,292	0,00046	-2,316	0,00340
Acsm2	protein_coding	ENSMUST00000084647	protein_coding	-1,399	0,00006	-3,448	0,00000
Acsm3	protein_coding	ENSMUST00000106527	protein_coding	-1,358	0,00007	-2,952	0,00000
Acsm3	protein_coding	ENSMUST00000063770	protein_coding	-1,557	0,00102	-2,486	0,00000
Acsm3	protein_coding	ENSMUST00000106529	protein_coding	-1,566	0,00702	-2,107	0,00001
Acsm5	protein_coding	ENSMUST00000207307	nonsense_mediated_decay	-1,090	0,00017	-2,222	0,00050
Acsm5	protein_coding	ENSMUST00000207440	protein_coding	-1,275	0,00029	-3,210	0,00019
Acss1	protein_coding	ENSMUST00000028944	protein_coding	-1,089	0,00007	-1,018	0,00002
Acss2	protein_coding	ENSMUST00000029135	protein_coding	-1,254	0,00011	-2,168	0,00000
Acss3	protein_coding	ENSMUST00000165067	protein_coding	-1,977	0,00007	-1,744	0,00490
Acy1	protein_coding	ENSMUST00000024031	protein_coding	-1,197	0,00007	-1,755	0,00000

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Acy1	protein_coding	ENSMUST00000190851	retained_intron	-1,294	0,00625	-2,090	0,00025
Acy3	protein_coding	ENSMUST00000054030	protein_coding	-1,687	0,00007	-2,680	0,00000
Adamts15	protein_coding	ENSMUST00000065112	protein_coding	-1,907	0,00006	-1,048	0,00024
Adhfe1	protein_coding	ENSMUST00000027044	nonsense_mediated_decay	-1,220	0,00009	-2,154	0,00000
Adhfe1	protein_coding	ENSMUST00000144177	protein_coding	-1,238	0,00067	-2,077	0,00000
Adk	protein_coding	ENSMUST00000045376	protein_coding	-2,055	0,00026	-1,404	0,00000
Adtrp	protein_coding	ENSMUST00000072012	protein_coding	-1,052	0,00044	-2,194	0,00000
Afap111	protein_coding	ENSMUST00000154876	nonsense_mediated_decay	-1,660	0,00015	-5,542	0,00056
Agxt2	protein_coding	ENSMUST00000110542	protein_coding	-1,166	0,00005	-2,199	0,00000
Agxt2	protein_coding	ENSMUST00000110541	protein_coding	-1,286	0,00324	-2,232	0,00147
Aim	lncRNA	ENSMUST00000160932	antisense	-1,304	0,00155	-2,009	0,00034
Ak4	protein_coding	ENSMUST00000106945	protein_coding	-2,124	0,00202	-4,062	0,00080
Akr1c14	protein_coding	ENSMUST00000041768	protein_coding	-1,171	0,00036	-2,880	0,00000
Akr1c14	protein_coding	ENSMUST00000118717	protein_coding	-1,318	0,00118	-3,192	0,00002
Akr1c21	protein_coding	ENSMUST000000223285	nonsense_mediated_decay	-1,034	0,00058	-2,657	0,00033
Akr1c21	protein_coding	ENSMUST000000021628	protein_coding	-1,267	0,00014	-4,349	0,00000
Akr1c21	protein_coding	ENSMUST00000137279	retained_intron	-1,487	0,00148	-3,441	0,00001
Akr1d1	protein_coding	ENSMUST00000040987	protein_coding	-1,106	0,00007	-2,413	0,00000
Akr1d1	protein_coding	ENSMUST00000123648	processed_transcript	-1,219	0,00063	-2,671	0,00012
Aldh7a1	protein_coding	ENSMUST00000174518	protein_coding	-1,044	0,00009	-1,461	0,00000
Aldob	protein_coding	ENSMUST00000144372	processed_transcript	-1,017	0,00802	-1,116	0,00143
Alms1-ps2	pseudogene	ENSMUST00000160606	processed_transcript	-1,289	0,00533	-4,015	0,00001
Amacr	protein_coding	ENSMUST00000070877	protein_coding	-1,378	0,00009	-2,299	0,00000
Amacr	protein_coding	ENSMUST000000228886	processed_transcript	-1,392	0,00006	-2,615	0,00000
Anks1b	protein_coding	ENSMUST00000182907	nonsense_mediated_decay	-3,821	0,00537	-4,760	0,00004
Ano3	protein_coding	ENSMUST00000099623	protein_coding	-1,284	0,00363	-1,058	0,00386
Aoah	protein_coding	ENSMUST00000021757	protein_coding	-1,437	0,00010	-1,655	0,00000
Apln	protein_coding	ENSMUST00000039026	protein_coding	-1,138	0,00067	-1,439	0,00028
Arhgef39	protein_coding	ENSMUST00000054538	protein_coding	-1,675	0,00039	-1,947	0,00007
Arhgef39	protein_coding	ENSMUST00000128754	processed_transcript	-1,870	0,00529	-2,073	0,00085
Arhgef39	protein_coding	ENSMUST00000125834	processed_transcript	-2,054	0,00723	-3,169	0,00409
Armh4	protein_coding	ENSMUST00000036972	protein_coding	-1,504	0,00133	-1,532	0,00000
Armh4	protein_coding	ENSMUST00000118129	protein_coding	-2,345	0,00059	-2,246	0,00038
Arsb	protein_coding	ENSMUST00000091403	protein_coding	-1,349	0,00006	-1,350	0,00000
Arsg	protein_coding	ENSMUST00000106697	protein_coding	-1,155	0,00146	-1,245	0,00002
Asb9	protein_coding	ENSMUST00000033756	protein_coding	-1,066	0,00075	-1,808	0,00045
Ascc1	protein_coding	ENSMUST00000050516	protein_coding	-1,559	0,00054	-1,843	0,00008
Aspa	protein_coding	ENSMUST00000021119	protein_coding	-1,280	0,00025	-2,744	0,00001
Aspa	protein_coding	ENSMUST00000132774	retained_intron	-1,416	0,00010	-1,613	0,00006
Aspa	protein_coding	ENSMUST00000141898	nonsense_mediated_decay	-1,578	0,00578	-2,372	0,00013
Aspdh	protein_coding	ENSMUST00000135624	protein_coding	-1,215	0,00203	-3,460	0,00009
Aspdh	protein_coding	ENSMUST00000146128	protein_coding	-1,552	0,00122	-4,804	0,00065
Aspdh	protein_coding	ENSMUST00000207591	TEC	-1,553	0,00033	-3,954	0,00018
Aspdh	protein_coding	ENSMUST00000035929	protein_coding	-1,570	0,00014	-4,303	0,00002
Aspdh	protein_coding	ENSMUST00000133801	retained_intron	-1,614	0,00057	-5,185	0,00003
Aspg	protein_coding	ENSMUST000000223457	retained_intron	-1,901	0,00866	-1,372	0,00620
Aspg	protein_coding	ENSMUST000000220719	nonsense_mediated_decay	-2,046	0,00034	-1,284	0,00003

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Aspg	protein_coding	ENSMUST00000079400	protein_coding	-2,197	0,00008	-1,912	0,00006
Aspg	protein_coding	ENSMUST000000221091	retained_intron	-2,595	0,00054	-1,639	0,00128
Ass1	protein_coding	ENSMUST000000102840	protein_coding	-1,513	0,00007	-2,467	0,00000
Ass1	protein_coding	ENSMUST000000192802	retained_intron	-1,653	0,00149	-1,669	0,00103
Ass1	protein_coding	ENSMUST000000130195	retained_intron	-1,783	0,00179	-1,954	0,00003
Atg10	protein_coding	ENSMUST00000022119	protein_coding	-2,173	0,00062	-1,741	0,00014
Atp11a	protein_coding	ENSMUST00000033818	protein_coding	-1,266	0,00006	-1,164	0,00051
Atp11a	protein_coding	ENSMUST000000133338	nonsense_mediated_decay	-1,279	0,00147	-1,564	0,00005
Atp11a	protein_coding	ENSMUST000000131804	protein_coding	-1,331	0,00043	-1,607	0,00005
Atp11a	protein_coding	ENSMUST000000139545	retained_intron	-1,420	0,00162	-1,447	0,00002
Atp11a	protein_coding	ENSMUST000000143359	protein_coding	-1,454	0,00043	-1,163	0,00011
Atp11a	protein_coding	ENSMUST000000132974	protein_coding	-1,716	0,00011	-2,835	0,00001
Atp11a	protein_coding	ENSMUST000000091237	protein_coding	-2,490	0,00043	-2,434	0,00004
Atm11	protein_coding	ENSMUST00000077282	protein_coding	-2,494	0,00002	-1,163	0,00000
Auh	protein_coding	ENSMUST00000021913	protein_coding	-1,056	0,00040	-1,466	0,00000
B3galnt1	protein_coding	ENSMUST000000061826	protein_coding	-1,511	0,00009	-2,182	0,00004
B3gat2	protein_coding	ENSMUST000000063663	protein_coding	-1,127	0,00009	-1,284	0,00000
Bbox1	protein_coding	ENSMUST000000046233	protein_coding	-1,493	0,00619	-2,399	0,00065
BC024386	lincRNA	ENSMUST000000209496	lincRNA	-1,986	0,00051	-3,558	0,00009
BC024386	lincRNA	ENSMUST000000209817	lincRNA	-2,017	0,00213	-4,882	0,00019
BC024386	lincRNA	ENSMUST000000209842	lincRNA	-2,547	0,00562	-2,305	0,00354
BC024386	lincRNA	ENSMUST000000209406	retained_intron	-2,561	0,00005	-3,463	0,00000
BC024386	lincRNA	ENSMUST000000210567	retained_intron	-2,663	0,00022	-7,304	0,00012
BC049987	lincRNA	ENSMUST000000213859	lincRNA	-1,573	0,00069	-2,102	0,00003
BC049987	lincRNA	ENSMUST000000217474	lincRNA	-1,654	0,00071	-4,037	0,00012
Bcas3	protein_coding	ENSMUST000000154396	protein_coding	-2,093	0,00344	-1,876	0,00741
Bcat1	protein_coding	ENSMUST000000111742	protein_coding	-1,180	0,00252	-1,133	0,00052
Bcat1	protein_coding	ENSMUST000000123930	protein_coding	-1,306	0,00365	-1,223	0,00127
Bcat2	protein_coding	ENSMUST000000209204	protein_coding	-1,070	0,00080	-1,428	0,00050
Bcl11b	protein_coding	ENSMUST000000066060	protein_coding	-1,351	0,00014	-1,040	0,00006
Bdh1	protein_coding	ENSMUST000000089759	protein_coding	-1,065	0,00008	-3,683	0,00000
Bdh1	protein_coding	ENSMUST000000115227	protein_coding	-1,347	0,00006	-2,363	0,00000
C030005K06Rik	lincRNA	ENSMUST000000129164	processed_transcript	-1,813	0,00048	-3,558	0,00000
C1qtnf3	protein_coding	ENSMUST000000022853	protein_coding	-1,935	0,00025	-3,255	0,00000
C1qtnf3	protein_coding	ENSMUST000000110523	protein_coding	-2,192	0,00015	-4,771	0,00005
C330002G04Rik	lincRNA	ENSMUST000000181623	lincRNA	-2,095	0,00068	-1,618	0,00000
C330021F23Rik	protein_coding	ENSMUST000000136592	protein_coding	-1,143	0,00355	-1,919	0,00295
Cabco1	protein_coding	ENSMUST00000020103	protein_coding	-1,603	0,00024	-2,203	0,00008
Camkmt	protein_coding	ENSMUST000000095188	protein_coding	-2,245	0,00060	-1,647	0,00006
Car12	protein_coding	ENSMUST000000152011	retained_intron	-1,108	0,00480	-1,117	0,00024
Car12	protein_coding	ENSMUST000000071889	protein_coding	-1,183	0,00020	-1,109	0,00001
Car14	protein_coding	ENSMUST000000149202	retained_intron	-1,660	0,00060	-1,946	0,00002
Car14	protein_coding	ENSMUST000000036181	protein_coding	-1,704	0,00011	-3,330	0,00000
Car4	protein_coding	ENSMUST000000103194	protein_coding	-1,165	0,00078	-2,973	0,00000
Car5a	protein_coding	ENSMUST000000057653	protein_coding	-1,246	0,00156	-1,580	0,00006
Car9	protein_coding	ENSMUST000000030183	protein_coding	-1,667	0,00060	-1,561	0,00002
Ccdc162	protein_coding	ENSMUST000000189488	protein_coding	-1,959	0,00573	-1,013	0,00225

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Ccdc91	protein_coding	ENSMUST00000032441	protein_coding	-1,084	0,00132	-1,197	0,00001
Cckar	protein_coding	ENSMUST00000031093	protein_coding	-1,042	0,00032	-2,033	0,00006
Ccser1	protein_coding	ENSMUST00000134710	protein_coding	-3,048	0,00062	-1,417	0,00009
Ccser2	protein_coding	ENSMUST00000182042	protein_coding	-1,027	0,00384	-1,709	0,00000
Cdh4	protein_coding	ENSMUST00000000314	protein_coding	-2,592	0,00367	-1,331	0,00005
Ces1e	protein_coding	ENSMUST00000034173	protein_coding	-1,574	0,00005	-3,738	0,00000
Ces1e	protein_coding	ENSMUST00000176282	protein_coding	-1,679	0,00038	-4,243	0,00009
Ces1e	protein_coding	ENSMUST00000175983	retained_intron	-1,774	0,00170	-3,677	0,00009
Ces1f	protein_coding	ENSMUST00000140026	nonsense_mediated_decay	-1,501	0,00074	-1,310	0,00519
Chchd6	protein_coding	ENSMUST00000032172	protein_coding	-1,888	0,00009	-1,503	0,00001
Chpt1	protein_coding	ENSMUST00000117579	protein_coding	-1,059	0,00389	-1,079	0,00006
Chpt1	protein_coding	ENSMUST00000117440	protein_coding	-1,212	0,00036	-1,583	0,00000
Chrd11	protein_coding	ENSMUST00000074660	protein_coding	-1,419	0,00170	-1,180	0,00343
Chma4	protein_coding	ENSMUST00000067120	protein_coding	-1,087	0,00010	-2,852	0,00000
Chma4	protein_coding	ENSMUST00000124400	protein_coding	-1,256	0,00694	-3,697	0,00153
Chst7	protein_coding	ENSMUST00000044138	protein_coding	-1,074	0,00126	-1,400	0,00000
Clec2h	protein_coding	ENSMUST00000032518	protein_coding	-1,807	0,00010	-2,863	0,00000
Clybl	protein_coding	ENSMUST00000026625	protein_coding	-1,675	0,00005	-1,522	0,00000
Cmah	protein_coding	ENSMUST00000167746	protein_coding	-1,319	0,00170	-1,987	0,00766
Cmss1	protein_coding	ENSMUST00000114371	protein_coding	-1,755	0,00022	-1,854	0,00005
Cndp1	protein_coding	ENSMUST00000130579	retained_intron	-1,756	0,00107	-1,249	0,00009
Cndp1	protein_coding	ENSMUST00000070139	protein_coding	-2,187	0,00005	-4,109	0,00000
Cndp2	protein_coding	ENSMUST00000025546	protein_coding	-1,306	0,00138	-2,080	0,00007
Col19a1	protein_coding	ENSMUST00000144297	processed_transcript	-2,300	0,00361	-3,189	0,00023
Commd1	protein_coding	ENSMUST00000093270	protein_coding	-1,475	0,00026	-1,006	0,00002
Coq8a	protein_coding	ENSMUST00000161300	protein_coding	-1,110	0,00079	-1,309	0,00004
Cox16	protein_coding	ENSMUST00000110340	protein_coding	-1,015	0,00094	-1,247	0,00020
Cpeb3	protein_coding	ENSMUST00000124158	protein_coding	-1,058	0,00242	-1,914	0,00000
Cpq	protein_coding	ENSMUST00000042167	protein_coding	-1,326	0,00006	-2,401	0,00000
Cry11	protein_coding	ENSMUST00000022517	protein_coding	-1,110	0,00009	-1,201	0,00000
Csgalnact1	protein_coding	ENSMUST00000078350	protein_coding	-1,635	0,00010	-1,626	0,00085
Ctnna2	protein_coding	ENSMUST00000162273	protein_coding	-1,448	0,00178	-1,768	0,00017
Ctnna2	protein_coding	ENSMUST00000161846	protein_coding	-4,717	0,00168	-4,475	0,00072
Cubn	protein_coding	ENSMUST00000091436	protein_coding	-1,658	0,00020	-1,759	0,00000
Cux1	protein_coding	ENSMUST00000176216	protein_coding	-1,326	0,00685	-1,763	0,00103
Cwh43	protein_coding	ENSMUST00000031040	protein_coding	-1,426	0,00357	-1,995	0,00000
Cym	protein_coding	ENSMUST00000029504	protein_coding	-1,280	0,00782	-2,744	0,00032
Cyp2d12	protein_coding	ENSMUST00000068861	protein_coding	-1,308	0,00849	-4,807	0,00000
Cyp2d9	protein_coding	ENSMUST00000229181	retained_intron	-1,039	0,00546	-2,292	0,00007
Cyp2d9	protein_coding	ENSMUST00000089129	protein_coding	-1,154	0,00062	-1,966	0,00001
Cyp2e1	protein_coding	ENSMUST00000209253	protein_coding	-1,528	0,00238	-3,058	0,00002
Cyp2e1	protein_coding	ENSMUST00000026552	protein_coding	-1,779	0,00007	-3,390	0,00002
Cyp2j11	protein_coding	ENSMUST00000015368	protein_coding	-1,312	0,00009	-3,183	0,00001
Cyp2j13	protein_coding	ENSMUST00000132513	protein_coding	-1,343	0,00103	-1,961	0,00027
Cyp2j5	protein_coding	ENSMUST00000030299	protein_coding	-1,048	0,00025	-2,643	0,00001
Cyp2j7	protein_coding	ENSMUST00000162514	protein_coding	-2,138	0,00312	-2,507	0,00001
Cyp2j9	protein_coding	ENSMUST00000055693	protein_coding	-1,069	0,00236	-1,081	0,00002

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Cyp2j9	protein_coding	ENSMUST00000126509	retained_intron	-1,387	0,00438	-1,131	0,00060
Cyp46a1	protein_coding	ENSMUST00000021684	protein_coding	-1,024	0,00864	-1,078	0,00074
Cyp4b1	protein_coding	ENSMUST00000102707	protein_coding	-1,179	0,00006	-1,706	0,00001
Cyp4b1	protein_coding	ENSMUST00000145841	protein_coding	-1,272	0,00085	-1,384	0,00007
Cyp4b1	protein_coding	ENSMUST00000135569	nonsense_mediated_decay	-1,463	0,00019	-1,368	0,00009
Cyp4b1	protein_coding	ENSMUST00000143237	retained_intron	-1,539	0,00025	-1,423	0,00114
Cyp51	protein_coding	ENSMUST00000001507	protein_coding	-1,643	0,00004	-2,378	0,00000
Cyp7b1	protein_coding	ENSMUST00000035625	protein_coding	-1,423	0,00045	-2,101	0,00002
D230017M19Rik	lincRNA	ENSMUST00000181973	lincRNA	-1,140	0,00876	-1,584	0,00026
D630023O14Rik	TEC	ENSMUST00000191913	TEC	-1,124	0,00563	-1,392	0,00008
D630024D03Rik	lincRNA	ENSMUST00000126265	lincRNA	-1,344	0,00074	-1,225	0,00020
D630029K05Rik	lincRNA	ENSMUST00000033651	lincRNA	-1,552	0,00008	-2,915	0,00000
D630029K05Rik	lincRNA	ENSMUST00000176050	lincRNA	-1,663	0,00266	-2,388	0,00042
D630029K05Rik	lincRNA	ENSMUST00000176455	lincRNA	-2,148	0,00009	-2,882	0,00000
Dbp	protein_coding	ENSMUST00000080885	protein_coding	-1,067	0,00238	-1,857	0,00001
Dhfr	protein_coding	ENSMUST00000022218	protein_coding	-1,228	0,00007	-1,548	0,00000
Dhtkd1	protein_coding	ENSMUST00000095147	protein_coding	-1,142	0,00031	-1,611	0,00000
Dmgdh	protein_coding	ENSMUST00000048001	protein_coding	-1,041	0,00007	-2,153	0,00000
Dnah5	protein_coding	ENSMUST00000067048	protein_coding	-2,340	0,00402	-3,200	0,00005
Dnah7b	protein_coding	ENSMUST00000069293	protein_coding	-1,753	0,00050	-1,658	0,00001
Dnah7c	protein_coding	ENSMUST00000189749	protein_coding	-2,170	0,00264	-2,822	0,00003
Dnajc12	protein_coding	ENSMUST00000043317	protein_coding	-1,120	0,00013	-2,069	0,00000
Dnajc12	protein_coding	ENSMUST00000129625	protein_coding	-1,326	0,00012	-2,586	0,00000
Dnajc12	protein_coding	ENSMUST00000136832	retained_intron	-1,511	0,00016	-2,658	0,00003
Dnase1	protein_coding	ENSMUST00000137748	protein_coding	-1,854	0,00031	-3,552	0,00005
Dnase1	protein_coding	ENSMUST00000125961	retained_intron	-1,926	0,00076	-3,023	0,00004
Dnase1	protein_coding	ENSMUST00000120009	protein_coding	-2,055	0,00006	-3,593	0,00001
Dnase1	protein_coding	ENSMUST00000144792	retained_intron	-2,067	0,00030	-2,031	0,00002
Dnase1	protein_coding	ENSMUST00000177337	nonsense_mediated_decay	-2,094	0,00039	-5,072	0,00093
Dnase1	protein_coding	ENSMUST00000157044	protein_coding	-2,415	0,00006	-3,622	0,00001
Dnase1	protein_coding	ENSMUST00000006136	protein_coding	-2,781	0,00005	-6,997	0,00250
Dnase1	protein_coding	ENSMUST00000132466	retained_intron	-4,800	0,00298	-4,070	0,00049
Dpyd	protein_coding	ENSMUST00000039177	protein_coding	-1,825	0,00011	-2,946	0,00000
Dtd1	protein_coding	ENSMUST00000028917	protein_coding	-1,526	0,00022	-1,245	0,00003
Eci3	protein_coding	ENSMUST00000164155	protein_coding	-1,103	0,00082	-3,293	0,00005
Eci3	protein_coding	ENSMUST00000021853	protein_coding	-1,207	0,00009	-3,260	0,00000
Eddm3b	protein_coding	ENSMUST00000100645	protein_coding	-2,145	0,00049	-2,316	0,00562
Efcab2	protein_coding	ENSMUST00000027775	protein_coding	-1,126	0,00028	-1,661	0,00003
Egf	protein_coding	ENSMUST00000197713	nonsense_mediated_decay	-1,514	0,00022	-2,732	0,00000
Egf	protein_coding	ENSMUST00000199272	retained_intron	-1,865	0,00016	-3,050	0,00000
Egf	protein_coding	ENSMUST00000029653	protein_coding	-1,937	0,00004	-4,316	0,00000
Egf	protein_coding	ENSMUST00000199615	protein_coding	-2,385	0,00201	-3,631	0,00002
Egf	protein_coding	ENSMUST00000197250	retained_intron	-2,496	0,00007	-2,200	0,00000
Eml6	protein_coding	ENSMUST00000058902	protein_coding	-1,818	0,00308	-1,147	0,00003
Enox1	protein_coding	ENSMUST00000227662	protein_coding	-4,033	0,00617	-3,374	0,00147
Enpp2	protein_coding	ENSMUST00000226339	protein_coding	-1,423	0,00466	-3,249	0,00087
Eny2	protein_coding	ENSMUST00000060652	protein_coding	-1,193	0,00011	-1,364	0,00011

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Epm2a	protein_coding	ENSMUST00000069106	protein_coding	-1,695	0,00006	-1,210	0,00002
Erc2	protein_coding	ENSMUST00000209800	retained_intron	-6,502	0,00006	-3,630	0,00039
Erc2	protein_coding	ENSMUST00000210924	protein_coding	-6,696	0,00005	-5,729	0,00000
Ero11b	protein_coding	ENSMUST00000220811	protein_coding	-1,178	0,00083	-1,235	0,00000
Esr1	protein_coding	ENSMUST00000105590	protein_coding	-2,280	0,00007	-2,135	0,00003
Esrsg	protein_coding	ENSMUST00000027906	protein_coding	-1,639	0,00083	-1,264	0,00139
Etnpp1	protein_coding	ENSMUST00000072271	protein_coding	-1,213	0,00025	-2,025	0,00003
F13b	protein_coding	ENSMUST00000027615	protein_coding	-1,436	0,00007	-3,290	0,00000
F8	protein_coding	ENSMUST00000114085	protein_coding	-1,074	0,00071	-2,015	0,00009
Fah	protein_coding	ENSMUST00000032865	protein_coding	-1,035	0,00014	-1,241	0,00000
Fam107a	protein_coding	ENSMUST00000036070	protein_coding	-1,261	0,00009	-2,352	0,00003
Fam78a	protein_coding	ENSMUST00000056406	protein_coding	-1,040	0,00684	-1,938	0,00030
Fars2	protein_coding	ENSMUST00000021857	protein_coding	-2,333	0,00082	-1,153	0,00034
Fars2	protein_coding	ENSMUST00000224241	protein_coding	-2,421	0,00015	-1,375	0,00016
Fars2	protein_coding	ENSMUST00000224611	protein_coding	-5,402	0,00060	-1,151	0,00876
Fbln7	protein_coding	ENSMUST00000028864	protein_coding	-1,414	0,00012	-1,398	0,00013
Fbp1	protein_coding	ENSMUST00000092888	protein_coding	-1,335	0,00016	-2,101	0,00000
Fcamr	protein_coding	ENSMUST00000027670	protein_coding	-1,726	0,00010	-2,674	0,00018
Fcamr	protein_coding	ENSMUST00000112477	protein_coding	-2,546	0,00057	-2,923	0,00028
Fggy	protein_coding	ENSMUST00000043335	protein_coding	-2,065	0,00115	-2,168	0,00025
Fggy	protein_coding	ENSMUST00000079223	protein_coding	-2,282	0,00006	-1,970	0,00001
Fggy	protein_coding	ENSMUST00000141248	processed_transcript	-2,690	0,00023	-1,751	0,00004
Fggy	protein_coding	ENSMUST00000147783	processed_transcript	-3,077	0,00760	-2,224	0,00037
Fhit	protein_coding	ENSMUST00000161302	protein_coding	-4,376	0,00203	-3,188	0,00005
Fhit	protein_coding	ENSMUST00000162278	protein_coding	-5,313	0,00034	-3,640	0,00006
Fhod3	protein_coding	ENSMUST00000234834	protein_coding	-1,299	0,00077	-1,409	0,00006
Fmod	protein_coding	ENSMUST00000048183	protein_coding	-1,476	0,00041	-1,834	0,00004
Fn3k	protein_coding	ENSMUST00000103014	protein_coding	-1,153	0,00361	-2,095	0,00001
Fn3k	protein_coding	ENSMUST00000026175	protein_coding	-1,193	0,00010	-2,731	0,00003
Fras1	protein_coding	ENSMUST00000036019	protein_coding	-1,868	0,00005	-1,618	0,00009
Frmpd3	protein_coding	ENSMUST00000208130	protein_coding	-1,374	0,00319	-1,092	0,00384
Fut9	protein_coding	ENSMUST00000084770	protein_coding	-1,448	0,00029	-2,353	0,00000
Fxyd2	protein_coding	ENSMUST00000217549	processed_transcript	-1,573	0,00185	-1,304	0,00001
Fxyd2	protein_coding	ENSMUST00000213850	protein_coding	-1,632	0,00017	-2,148	0,00002
Fxyd2	protein_coding	ENSMUST00000214680	processed_transcript	-2,125	0,00021	-3,207	0,00001
G6pc	protein_coding	ENSMUST00000019469	protein_coding	-1,701	0,00095	-2,908	0,00000
Gabrb3	protein_coding	ENSMUST00000039697	protein_coding	-1,533	0,00006	-1,974	0,00024
Galnt11	protein_coding	ENSMUST00000114950	protein_coding	-1,975	0,00027	-3,770	0,00000
Galnt14	protein_coding	ENSMUST00000112591	protein_coding	-1,237	0,00041	-1,155	0,00047
Galnt14	protein_coding	ENSMUST00000024858	protein_coding	-1,475	0,00985	-1,468	0,00007
Galnt14	protein_coding	ENSMUST00000232785	processed_transcript	-1,525	0,00131	-1,535	0,00135
Gamt	protein_coding	ENSMUST00000105363	protein_coding	-1,114	0,00074	-1,956	0,00000
Gatb	protein_coding	ENSMUST00000029726	processed_transcript	-1,004	0,00012	-1,388	0,00000
Gatb	protein_coding	ENSMUST00000127348	protein_coding	-1,084	0,00026	-1,235	0,00000
Gatm	protein_coding	ENSMUST00000140808	retained_intron	-1,775	0,00035	-2,092	0,00007
Gatm	protein_coding	ENSMUST00000028624	protein_coding	-2,612	0,00003	-3,170	0,00000
Gcdh	protein_coding	ENSMUST00000139180	retained_intron	-1,216	0,00091	-1,435	0,00005

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Ggnbp1	protein_coding	ENSMUST0000053683	protein_coding	-1,757	0,00419	-1,556	0,00217
Ggt1	protein_coding	ENSMUST00000125770	protein_coding	-1,438	0,00012	-3,222	0,00007
Ghr	protein_coding	ENSMUST00000110698	protein_coding	-1,157	0,00010	-1,221	0,00000
Ghr	protein_coding	ENSMUST00000069451	protein_coding	-1,496	0,00084	-1,279	0,00004
Ghr	protein_coding	ENSMUST00000161770	protein_coding	-1,809	0,00032	-1,296	0,00001
Gk	protein_coding	ENSMUST00000156390	protein_coding	-1,040	0,00259	-1,896	0,00002
Gk	protein_coding	ENSMUST00000142152	protein_coding	-1,524	0,00083	-2,295	0,00000
Gk	protein_coding	ENSMUST00000113978	protein_coding	-1,834	0,00525	-2,006	0,00023
Gm10804	lncRNA	ENSMUST00000162565	lncRNA	-1,171	0,00016	-1,602	0,00001
Gm12678	lncRNA	ENSMUST00000154177	lncRNA	-1,446	0,00220	-3,599	0,00001
Gm15318	lncRNA	ENSMUST00000146320	antisense	-1,051	0,00042	-2,303	0,00014
Gm15318	lncRNA	ENSMUST00000139773	antisense	-1,219	0,00711	-1,941	0,00277
Gm15347	lncRNA	ENSMUST00000145372	antisense	-1,454	0,00545	-1,203	0,00013
Gm15348	lncRNA	ENSMUST00000151188	lncRNA	-1,104	0,00024	-1,677	0,00002
Gm15848	lncRNA	ENSMUST00000125056	antisense	-1,267	0,00022	-1,867	0,00182
Gm15848	lncRNA	ENSMUST00000136015	antisense	-2,160	0,00009	-2,488	0,00012
Gm15848	lncRNA	ENSMUST00000122968	antisense	-2,285	0,00042	-2,487	0,00023
Gm15883	lncRNA	ENSMUST00000153562	antisense	-1,555	0,00190	-1,307	0,00036
Gm17597	lncRNA	ENSMUST00000181220	antisense	-1,219	0,00244	-1,315	0,00090
Gm20449	protein_coding	ENSMUST00000094532	nonsense_mediated_decay	-1,231	0,00208	-1,046	0,00023
Gm23332	misc_RNA	ENSMUST00000157424	misc_RNA	-1,968	0,00216	-3,011	0,00013
Gm28023	pseudogene	ENSMUST00000230537	unprocessed_pseudogene	-1,425	0,00432	-3,728	0,00003
Gm28437	pseudogene	ENSMUST00000190277	unprocessed_pseudogene	-1,325	0,00779	-1,232	0,00000
Gm29012	lncRNA	ENSMUST00000187882	lncRNA	-2,863	0,00060	-3,782	0,00081
Gm31718	lncRNA	ENSMUST00000212639	lncRNA	-1,129	0,00103	-1,143	0,00589
Gm3435	protein_coding	ENSMUST00000097395	protein_coding	-4,770	0,00713	-5,396	0,00006
Gm34376	lncRNA	ENSMUST00000218529	lncRNA	-2,034	0,00156	-1,726	0,00435
Gm37260	TEC	ENSMUST00000193748	TEC	-2,151	0,00105	-1,515	0,00017
Gm4131	protein_coding	ENSMUST00000186010	protein_coding	-1,292	0,00079	-1,963	0,00008
Gm4208	lncRNA	ENSMUST00000189802	lncRNA	-3,570	0,00006	-2,187	0,00017
Gm42397	lncRNA	ENSMUST00000208808	lncRNA	-3,650	0,00815	-2,936	0,00000
Gm44000	lncRNA	ENSMUST00000205038	antisense	-2,342	0,00142	-1,885	0,00002
Gm44117	lncRNA	ENSMUST00000204442	lncRNA	-1,237	0,00500	-1,871	0,00330
Gm48197	lncRNA	ENSMUST00000213800	lncRNA	-4,471	0,00479	-1,332	0,00091
Gm48281	lncRNA	ENSMUST00000217707	lncRNA	-2,738	0,00760	-3,774	0,00041
Gm48702	lncRNA	ENSMUST00000214931	antisense	-1,418	0,00217	-2,540	0,00001
Gm49179	lncRNA	ENSMUST00000226752	antisense	-1,082	0,00298	-1,457	0,00009
Gm49542	lncRNA	ENSMUST00000227349	lncRNA	-1,986	0,00005	-2,436	0,00003
Gm5524	lncRNA	ENSMUST00000185409	lncRNA	-1,188	0,00113	-3,606	0,00032
Gm5717	pseudogene	ENSMUST00000198243	unprocessed_pseudogene	-1,345	0,00046	-2,969	0,00007
Gm6300	pseudogene	ENSMUST00000192393	unprocessed_pseudogene	-1,640	0,00849	-4,485	0,00002
Gm6999	lncRNA	ENSMUST00000232826	lncRNA	-2,192	0,00031	-2,939	0,00001
Gm6999	lncRNA	ENSMUST00000232869	lncRNA	-2,509	0,00111	-2,395	0,00002
Gm6999	lncRNA	ENSMUST00000227982	lncRNA	-2,699	0,00181	-2,362	0,00001
Gna14	protein_coding	ENSMUST00000025602	protein_coding	-1,645	0,00010	-1,580	0,00000
Gphn	protein_coding	ENSMUST00000110388	protein_coding	-2,379	0,00014	-1,333	0,00001
Gpm6a	protein_coding	ENSMUST00000033915	protein_coding	-1,821	0,00017	-1,218	0,00000

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Gpr137b	protein_coding	ENSMUST00000222958	nonsense_mediated_decay	-1,397	0,00502	-1,799	0,00004
Gpr137b-ps	pseudogene	ENSMUST00000223146	retained_intron	-1,048	0,00191	-1,839	0,00007
Gpx3	protein_coding	ENSMUST00000125094	protein_coding	-1,345	0,00289	-1,661	0,00951
Gys2	protein_coding	ENSMUST00000032371	protein_coding	-1,863	0,00006	-3,621	0,00004
Hao	protein_coding	ENSMUST00000234750	retained_intron	-1,071	0,00795	-1,421	0,00044
Hao	protein_coding	ENSMUST00000234214	protein_coding	-1,183	0,00018	-2,315	0,00002
Hao	protein_coding	ENSMUST00000000687	protein_coding	-1,199	0,00010	-2,448	0,00001
Hao	protein_coding	ENSMUST00000234863	nonsense_mediated_decay	-1,217	0,00313	-2,987	0,00006
Hba-a2	protein_coding	ENSMUST00000093207	protein_coding	-1,515	0,00948	-4,805	0,00095
Hhat	protein_coding	ENSMUST00000044190	protein_coding	-1,611	0,00130	-1,209	0,00074
Hlf	protein_coding	ENSMUST00000004051	protein_coding	-1,539	0,00065	-2,834	0,00001
Hnf4aos	lncRNA	ENSMUST00000141329	antisense	-2,033	0,00073	-2,732	0,00002
Hook1	protein_coding	ENSMUST00000140128	processed_transcript	-1,285	0,00266	-1,010	0,00117
Hpd	protein_coding	ENSMUST00000031398	protein_coding	-1,670	0,00033	-3,419	0,00000
Hpd	protein_coding	ENSMUST00000154713	protein_coding	-1,746	0,00257	-3,473	0,00162
Hpd	protein_coding	ENSMUST00000156539	retained_intron	-2,286	0,00063	-2,562	0,00015
Hrg	protein_coding	ENSMUST00000023590	protein_coding	-1,756	0,00286	-3,250	0,00010
Hsd3b2	protein_coding	ENSMUST00000107022	protein_coding	-1,785	0,00028	-2,474	0,00000
Hsd3b2	protein_coding	ENSMUST00000107021	protein_coding	-1,983	0,00007	-7,221	0,00001
Hykk	protein_coding	ENSMUST00000039742	protein_coding	-1,528	0,00010	-2,064	0,00000
Idh2	protein_coding	ENSMUST00000107384	protein_coding	-1,077	0,00008	-1,713	0,00000
Ido2	protein_coding	ENSMUST00000121992	protein_coding	-1,040	0,00246	-1,340	0,00250
Ildf2	protein_coding	ENSMUST00000195746	protein_coding	-1,117	0,00035	-1,961	0,00177
Ift122	protein_coding	ENSMUST00000038234	protein_coding	-1,056	0,00606	-1,728	0,00013
Igfbp1	protein_coding	ENSMUST00000020704	protein_coding	-1,967	0,00018	-1,243	0,00869
Immp21	protein_coding	ENSMUST00000134965	protein_coding	-3,815	0,00441	-1,972	0,00132
Inmt	protein_coding	ENSMUST00000003569	protein_coding	-1,577	0,00011	-4,768	0,00000
Inmt	protein_coding	ENSMUST00000152720	retained_intron	-2,297	0,00031	-4,353	0,00007
Invs	protein_coding	ENSMUST00000030029	protein_coding	-1,201	0,00026	-1,198	0,00016
Ipcefl	protein_coding	ENSMUST00000105617	protein_coding	-1,018	0,00031	-1,263	0,00000
Iqsec2	protein_coding	ENSMUST00000128724	processed_transcript	-1,179	0,00172	-1,844	0,00005
Iqsec2	protein_coding	ENSMUST00000146835	processed_transcript	-1,222	0,00123	-1,108	0,00013
Isoc2a	protein_coding	ENSMUST00000125249	protein_coding	-1,158	0,00005	-1,703	0,00000
Itih2	protein_coding	ENSMUST00000042290	protein_coding	-1,647	0,00026	-2,144	0,00000
Itih5	protein_coding	ENSMUST00000026886	protein_coding	-1,005	0,00027	-1,758	0,00000
Ivd	protein_coding	ENSMUST00000152191	processed_transcript	-1,094	0,00007	-1,457	0,00000
Ivd	protein_coding	ENSMUST00000150162	processed_transcript	-1,284	0,00105	-1,465	0,00000
Josd2	protein_coding	ENSMUST00000035844	protein_coding	-1,113	0,00819	-1,374	0,00857
Kap	protein_coding	ENSMUST00000203168	protein_coding	-1,184	0,00126	-2,584	0,00012
Kcng1	protein_coding	ENSMUST00000109191	protein_coding	-2,421	0,00013	-2,110	0,00107
Kcnj15	protein_coding	ENSMUST00000113854	protein_coding	-1,148	0,00118	-1,365	0,00180
Kcnma1	protein_coding	ENSMUST00000145596	protein_coding	-2,903	0,00336	-1,441	0,00015
Kl1	protein_coding	ENSMUST00000078856	protein_coding	-1,004	0,00007	-2,740	0,00000
Klfl2	protein_coding	ENSMUST00000228216	protein_coding	-2,332	0,00014	-2,231	0,00000
Klfl5	protein_coding	ENSMUST00000203607	protein_coding	-1,488	0,00329	-2,390	0,00058
Ksr2	protein_coding	ENSMUST00000180430	protein_coding	-2,428	0,00027	-2,207	0,00001
Lactb2	protein_coding	ENSMUST00000027071	protein_coding	-1,081	0,00018	-2,013	0,00000

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2

(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Lamp2	protein_coding	ENSMUST00000074913	protein_coding	-1,033	0,00029	-1,167	0,00001
Large1	protein_coding	ENSMUST00000004497	protein_coding	-2,184	0,00204	-1,647	0,00024
Lpl	protein_coding	ENSMUST00000015712	protein_coding	-1,361	0,00016	-1,808	0,00053
Lrat	protein_coding	ENSMUST00000029632	protein_coding	-2,357	0,00031	-1,849	0,00001
Lrp2	protein_coding	ENSMUST00000080953	protein_coding	-1,382	0,00023	-2,419	0,00000
Lrrc19	protein_coding	ENSMUST00000107101	protein_coding	-1,082	0,00550	-1,710	0,00041
Lrrk2	protein_coding	ENSMUST00000133743	processed_transcript	-1,020	0,00422	-1,737	0,00054
Lrrk2	protein_coding	ENSMUST00000140734	processed_transcript	-1,101	0,00282	-1,507	0,00012
Lrrk2	protein_coding	ENSMUST00000137657	retained_intron	-1,162	0,00480	-1,231	0,00005
Lrm1	protein_coding	ENSMUST00000049285	protein_coding	-1,571	0,00135	-1,315	0,00015
Lsamp	protein_coding	ENSMUST00000099761	protein_coding	-3,253	0,00632	-3,358	0,00039
MacroD2	protein_coding	ENSMUST00000110064	protein_coding	-1,498	0,00029	-2,625	0,00001
MacroD2	protein_coding	ENSMUST00000138786	processed_transcript	-1,954	0,00011	-2,911	0,00001
Malrd1	protein_coding	ENSMUST00000146205	protein_coding	-1,839	0,00016	-2,008	0,00000
Map3k15	protein_coding	ENSMUST00000033665	protein_coding	-1,385	0,00039	-1,218	0,00000
Mccc2	protein_coding	ENSMUST00000022148	protein_coding	-1,074	0,00023	-1,804	0,00000
Me3	protein_coding	ENSMUST00000032853	processed_transcript	-1,076	0,00144	-1,244	0,00031
Mep1b	protein_coding	ENSMUST00000082235	protein_coding	-1,928	0,00009	-4,616	0,00000
Mgam	protein_coding	ENSMUST00000202779	protein_coding	-1,682	0,00009	-1,971	0,00000
Mgam	protein_coding	ENSMUST00000071535	protein_coding	-1,737	0,00009	-2,042	0,00000
Mgam	protein_coding	ENSMUST00000201148	protein_coding	-1,737	0,00009	-2,042	0,00000
Mgat3	protein_coding	ENSMUST00000044970	protein_coding	-1,119	0,00014	-2,495	0,00000
Mia2	protein_coding	ENSMUST00000177370	protein_coding	-1,058	0,00232	-1,463	0,00005
Miox	protein_coding	ENSMUST00000160491	retained_intron	-1,207	0,00079	-2,827	0,00008
Miox	protein_coding	ENSMUST00000023282	protein_coding	-1,642	0,00006	-3,945	0,00000
Miox	protein_coding	ENSMUST00000161437	retained_intron	-1,928	0,00364	-2,017	0,00004
Mpv171	protein_coding	ENSMUST00000148966	nonsense_mediated_decay	-1,071	0,00924	-2,515	0,00011
Mpv171	protein_coding	ENSMUST00000128757	protein_coding	-1,210	0,00147	-1,319	0,00002
Mpv171	protein_coding	ENSMUST00000141971	protein_coding	-1,328	0,00229	-1,890	0,00075
Mpv171	protein_coding	ENSMUST00000023360	protein_coding	-1,355	0,00010	-2,099	0,00000
Msra	protein_coding	ENSMUST00000210428	protein_coding	-1,767	0,00006	-1,074	0,00000
Myo5a	protein_coding	ENSMUST00000123128	protein_coding	-1,031	0,00091	-1,195	0,00006
Myo5a	protein_coding	ENSMUST00000136604	processed_transcript	-1,292	0,00253	-5,665	0,00779
Nabp1	protein_coding	ENSMUST00000186684	protein_coding	-5,112	0,00321	-5,405	0,00087
Napsa	protein_coding	ENSMUST00000153124	retained_intron	-1,169	0,00123	-2,061	0,00000
Nat8f1	protein_coding	ENSMUST00000161198	protein_coding	-1,536	0,00036	-3,534	0,00000
Nat8f1	protein_coding	ENSMUST00000174369	protein_coding	-1,697	0,00115	-4,047	0,00430
Nat8f3	protein_coding	ENSMUST00000168531	protein_coding	-1,241	0,00533	-5,648	0,00032
Nat8f6	protein_coding	ENSMUST00000174143	protein_coding	-1,422	0,00359	-5,054	0,00000
Nat8f6	protein_coding	ENSMUST00000087656	protein_coding	-1,640	0,00141	-5,735	0,00000
Ncam2	protein_coding	ENSMUST00000067602	protein_coding	-2,397	0,00894	-1,207	0,00230
Nepn	protein_coding	ENSMUST000000219730	protein_coding	-1,404	0,00282	-2,732	0,00047
Neurl4	protein_coding	ENSMUST00000175761	retained_intron	-1,520	0,00085	-2,944	0,00043
Ngef	protein_coding	ENSMUST00000068681	protein_coding	-1,153	0,00052	-1,317	0,00000
Nlrp6	protein_coding	ENSMUST00000106045	protein_coding	-1,635	0,00010	-5,579	0,00732
Nlrp6	protein_coding	ENSMUST00000183845	protein_coding	-1,637	0,00016	-2,311	0,00000
Nnt	protein_coding	ENSMUST00000223268	protein_coding	-1,085	0,00142	-1,461	0,00000

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Nox4	protein_coding	ENSMUST00000126887	nonsense_mediated_decay	-1,028	0,00390	-2,701	0,00423
Nsg2	protein_coding	ENSMUST00000020537	protein_coding	-1,340	0,00432	-3,417	0,00295
Nubpl	protein_coding	ENSMUST00000162111	processed_transcript	-1,619	0,00282	-1,220	0,00209
Nubpl	protein_coding	ENSMUST00000040090	protein_coding	-1,960	0,00006	-1,238	0,00001
Nudt19	protein_coding	ENSMUST00000142930	processed_transcript	-1,071	0,00132	-2,967	0,00001
Nudt19	protein_coding	ENSMUST00000123469	retained_intron	-1,430	0,00342	-1,914	0,00023
Nxph4	protein_coding	ENSMUST00000095266	protein_coding	-1,899	0,00078	-2,432	0,00476
Osbpl6	protein_coding	ENSMUST00000077972	protein_coding	-1,158	0,00102	-7,800	0,00000
Oxct1	protein_coding	ENSMUST00000110690	protein_coding	-1,084	0,00006	-2,524	0,00000
Pacrg	protein_coding	ENSMUST00000041463	protein_coding	-2,534	0,00049	-1,839	0,00030
Pah	protein_coding	ENSMUST00000020241	protein_coding	-1,187	0,00006	-3,203	0,00000
Pank1	protein_coding	ENSMUST00000112460	protein_coding	-1,537	0,00022	-1,107	0,00004
Pard3	protein_coding	ENSMUST00000161348	protein_coding	-1,463	0,00205	-1,116	0,00041
Pbld1	protein_coding	ENSMUST00000219829	protein_coding	-1,225	0,00147	-2,158	0,00004
Pbld1	protein_coding	ENSMUST00000178684	protein_coding	-1,477	0,00006	-2,942	0,00000
Pcca	protein_coding	ENSMUST00000038374	protein_coding	-1,833	0,00004	-1,706	0,00005
Pck1	protein_coding	ENSMUST00000151269	retained_intron	-2,132	0,00069	-1,843	0,00004
Pck1	protein_coding	ENSMUST00000136885	retained_intron	-2,478	0,00632	-1,457	0,00003
Pclo	protein_coding	ENSMUST00000182915	protein_coding	-1,605	0,00156	-3,418	0,00703
Pcx	protein_coding	ENSMUST00000113825	protein_coding	-2,761	0,00424	-2,938	0,00009
Pcyt2	protein_coding	ENSMUST00000124504	retained_intron	-1,030	0,00598	-1,728	0,00061
Pcyt2	protein_coding	ENSMUST00000106187	retained_intron	-1,049	0,00045	-2,051	0,00005
Pde11a	protein_coding	ENSMUST00000099992	protein_coding	-2,077	0,00549	-2,868	0,00172
Pde6a	protein_coding	ENSMUST00000025468	protein_coding	-2,519	0,00008	-4,328	0,00000
Pde9a	protein_coding	ENSMUST00000047168	protein_coding	-1,271	0,00188	-1,031	0,00396
Pdgfd	protein_coding	ENSMUST00000058692	protein_coding	-1,368	0,00095	-1,032	0,00028
Pdss2	protein_coding	ENSMUST00000095725	protein_coding	-1,666	0,00010	-1,003	0,00012
Pdzd2	protein_coding	ENSMUST00000075317	protein_coding	-1,448	0,00035	-1,318	0,00001
Pdzk1	protein_coding	ENSMUST00000128789	protein_coding	-1,150	0,00841	-1,633	0,00003
Pdzm3	protein_coding	ENSMUST00000075994	protein_coding	-1,509	0,00411	-1,200	0,00000
Pecr	protein_coding	ENSMUST00000027381	protein_coding	-1,020	0,00054	-2,465	0,00000
Pepd	protein_coding	ENSMUST00000075068	protein_coding	-1,094	0,00011	-1,611	0,00000
Pi4ka	protein_coding	ENSMUST00000036161	protein_coding	-1,182	0,00008	-1,202	0,00001
Pnpla7	protein_coding	ENSMUST00000045295	protein_coding	-1,166	0,00065	-1,174	0,00003
Pola1	protein_coding	ENSMUST00000127718	retained_intron	-1,264	0,00208	-1,131	0,00032
Polq	protein_coding	ENSMUST00000182717	retained_intron	-2,010	0,00399	-1,449	0,00054
Polr3e	protein_coding	ENSMUST00000033173	protein_coding	-1,693	0,00035	-1,125	0,00482
Ppm1k	protein_coding	ENSMUST00000204607	protein_coding	-3,310	0,00140	-2,443	0,00000
Ppm1l	protein_coding	ENSMUST00000029355	protein_coding	-1,235	0,00014	-1,540	0,00001
Ppt2	protein_coding	ENSMUST00000167097	protein_coding	-1,109	0,00018	-1,713	0,00002
Prkecz	protein_coding	ENSMUST00000103178	protein_coding	-1,247	0,00011	-1,818	0,00000
Prkn	protein_coding	ENSMUST00000191124	protein_coding	-5,799	0,00042	-6,244	0,00000
Prodh	protein_coding	ENSMUST00000232384	retained_intron	-1,103	0,00987	-2,078	0,00000
Prodh2	protein_coding	ENSMUST00000142514	processed_transcript	-1,088	0,00076	-1,356	0,00001
Prodh2	protein_coding	ENSMUST00000058280	protein_coding	-1,261	0,00034	-3,354	0,00012
Psmc3	protein_coding	ENSMUST00000185715	protein_coding	-6,370	0,00087	-6,333	0,00001
Pth1r	protein_coding	ENSMUST00000199791	protein_coding	-1,178	0,00235	-1,942	0,00115

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2

(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Pth1r	protein_coding	ENSMUST00000198865	protein_coding	-1,206	0,00019	-2,621	0,00000
Ptprd	protein_coding	ENSMUST00000107289	protein_coding	-1,832	0,00050	-1,642	0,00000
Ptprd	protein_coding	ENSMUST00000102834	protein_coding	-1,955	0,00007	-1,702	0,00002
Ptprg	protein_coding	ENSMUST00000022264	protein_coding	-1,895	0,00004	-1,327	0,00000
Pvalb	protein_coding	ENSMUST00000005860	protein_coding	-1,570	0,00013	-4,514	0,00114
Pxmp2	protein_coding	ENSMUST000000031472	protein_coding	-1,154	0,00020	-1,791	0,00000
Pzp	protein_coding	ENSMUST00000112132	protein_coding	-1,953	0,00002	-3,707	0,00001
R3hdml	protein_coding	ENSMUST00000109416	protein_coding	-1,507	0,00403	-1,493	0,00724
Rab11fip3	protein_coding	ENSMUST00000118828	protein_coding	-1,035	0,00008	-1,982	0,00000
Ralgps2	protein_coding	ENSMUST00000172057	protein_coding	-1,038	0,00016	-1,421	0,00002
Rbfox1	protein_coding	ENSMUST00000056416	protein_coding	-3,741	0,00395	-3,067	0,00004
Reps2	protein_coding	ENSMUST00000154424	protein_coding	-1,001	0,00035	-1,922	0,00001
Rgl1	protein_coding	ENSMUST00000111859	protein_coding	-1,201	0,00054	-1,115	0,00020
Rgs6	protein_coding	ENSMUST00000186458	protein_coding	-2,097	0,00014	-1,116	0,00006
Rims2	protein_coding	ENSMUST00000226410	protein_coding	-1,287	0,00153	-1,018	0,00114
Rnls	pseudogene	ENSMUST00000096114	protein_coding	-2,409	0,00861	-1,809	0,00762
Rundc3b	protein_coding	ENSMUST00000047485	protein_coding	-1,900	0,00467	-1,736	0,00411
Ryr2	protein_coding	ENSMUST00000021750	protein_coding	-2,202	0,00443	-2,751	0,00004
Sardh	protein_coding	ENSMUST00000102886	protein_coding	-1,035	0,00006	-1,737	0,00001
Scfd2	protein_coding	ENSMUST00000146015	retained_intron	-1,847	0,00044	-2,701	0,00090
Scfd2	protein_coding	ENSMUST00000072857	protein_coding	-2,260	0,00016	-1,193	0,00032
Scfd2	protein_coding	ENSMUST00000113542	protein_coding	-5,076	0,00137	-4,426	0,00127
Scp2-ps2	pseudogene	ENSMUST00000216438	processed_transcript	-1,747	0,00030	-2,791	0,00999
Sdc2	protein_coding	ENSMUST00000022871	protein_coding	-1,262	0,00006	-1,318	0,00001
Sec14l3	protein_coding	ENSMUST00000068322	protein_coding	-1,132	0,00672	-1,464	0,00020
Serpina1f	protein_coding	ENSMUST00000117053	protein_coding	-1,132	0,00492	-1,212	0,00105
Serpina1f	protein_coding	ENSMUST00000118101	protein_coding	-1,139	0,00069	-1,162	0,00128
Serpinf2	protein_coding	ENSMUST00000128330	protein_coding	-1,216	0,00083	-2,719	0,00007
Serpinf2	protein_coding	ENSMUST00000043696	protein_coding	-1,391	0,00006	-2,689	0,00000
Sfxn5	protein_coding	ENSMUST00000045846	protein_coding	-1,563	0,00005	-1,512	0,00000
Shank2	protein_coding	ENSMUST00000105902	protein_coding	-1,368	0,00276	-1,274	0,00020
Shmt2	protein_coding	ENSMUST00000217682	retained_intron	-1,064	0,00100	-1,657	0,00003
Slc10a5	protein_coding	ENSMUST00000078748	protein_coding	-1,265	0,00013	-2,421	0,00000
Slc12a1	protein_coding	ENSMUST00000110494	protein_coding	-1,080	0,00012	-2,746	0,00017
Slc12a3	protein_coding	ENSMUST00000212041	retained_intron	-1,337	0,00991	-4,879	0,00001
Slc12a3	protein_coding	ENSMUST00000212134	protein_coding	-1,727	0,00065	-2,459	0,00000
Slc12a3	protein_coding	ENSMUST00000211905	protein_coding	-1,785	0,00051	-4,195	0,00067
Slc12a3	protein_coding	ENSMUST00000034218	protein_coding	-1,917	0,00006	-3,339	0,00000
Slc12a3	protein_coding	ENSMUST00000212915	retained_intron	-2,238	0,00006	-4,167	0,00000
Slc12a3	protein_coding	ENSMUST00000212632	retained_intron	-2,302	0,00004	-3,733	0,00000
Slc12a3	protein_coding	ENSMUST00000211942	retained_intron	-2,322	0,00002	-4,080	0,00000
Slc13a2	protein_coding	ENSMUST00000001122	protein_coding	-1,539	0,00006	-1,815	0,00147
Slc13a3	protein_coding	ENSMUST00000109279	protein_coding	-1,448	0,00031	-1,820	0,00001
Slc13a3	protein_coding	ENSMUST00000029208	protein_coding	-1,466	0,00019	-1,943	0,00004
Slc16a2	protein_coding	ENSMUST00000042664	protein_coding	-1,062	0,00011	-1,384	0,00000
Slc17a3	protein_coding	ENSMUST00000039721	protein_coding	-1,271	0,00015	-2,636	0,00005
Slc17a3	protein_coding	ENSMUST00000166467	protein_coding	-1,411	0,00427	-2,317	0,00099

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Slc18a1	protein_coding	ENSMUST00000037478	protein_coding	-1,033	0,00008	-1,745	0,00000
Slc22a18	protein_coding	ENSMUST00000105917	protein_coding	-1,385	0,00039	-3,795	0,00002
Slc22a22	protein_coding	ENSMUST00000022995	protein_coding	-1,534	0,00020	-4,608	0,00000
Slc22a28	protein_coding	ENSMUST00000065651	protein_coding	-2,006	0,00014	-4,885	0,00000
Slc22a30	protein_coding	ENSMUST00000064507	protein_coding	-1,433	0,00909	-4,186	0,00190
Slc22a30	protein_coding	ENSMUST00000096269	protein_coding	-1,887	0,00022	-4,099	0,00000
Slc22a30	protein_coding	ENSMUST00000120540	protein_coding	-2,514	0,00016	-4,913	0,00000
Slc22a8	protein_coding	ENSMUST00000010251	protein_coding	-1,205	0,00079	-3,650	0,00000
Slc25a13	protein_coding	ENSMUST00000015256	protein_coding	-1,407	0,00054	-1,389	0,00009
Slc25a21	protein_coding	ENSMUST00000044634	protein_coding	-3,421	0,00028	-1,547	0,00001
Slc25a21	protein_coding	ENSMUST00000217690	protein_coding	-6,368	0,00011	-1,487	0,00005
Slc27a2	protein_coding	ENSMUST00000061491	protein_coding	-1,622	0,00017	-1,595	0,00000
Slc27a2	protein_coding	ENSMUST00000141482	protein_coding	-1,643	0,00031	-1,568	0,00000
Slc2a5	protein_coding	ENSMUST00000030826	protein_coding	-1,277	0,00019	-2,484	0,00031
Slc34a1	protein_coding	ENSMUST00000224043	retained_intron	-1,404	0,00080	-4,137	0,00000
Slc34a1	protein_coding	ENSMUST00000057167	protein_coding	-1,472	0,00005	-4,538	0,00000
Slc34a1	protein_coding	ENSMUST00000224925	protein_coding	-1,612	0,00063	-4,827	0,00001
Slc34a1	protein_coding	ENSMUST00000225259	protein_coding	-1,644	0,00006	-4,168	0,00000
Slc34a1	protein_coding	ENSMUST00000223954	retained_intron	-1,874	0,00057	-4,987	0,00001
Slc34a1	protein_coding	ENSMUST00000225538	retained_intron	-1,896	0,00088	-4,859	0,00000
Slc35f1	protein_coding	ENSMUST00000105473	protein_coding	-1,481	0,00130	-1,704	0,00001
Slc37a4	protein_coding	ENSMUST00000165839	protein_coding	-1,227	0,00056	-2,039	0,00000
Slc37a4	protein_coding	ENSMUST00000215001	protein_coding	-1,570	0,00963	-1,692	0,00021
Slc46a3	protein_coding	ENSMUST00000031655	protein_coding	-1,050	0,00768	-2,350	0,00000
Slc47a1	protein_coding	ENSMUST00000147583	processed_transcript	-1,142	0,00064	-1,562	0,00001
Slc47a1	protein_coding	ENSMUST00000131723	nonsense_mediated_decay	-1,189	0,00154	-1,497	0,00000
Slc4a4	protein_coding	ENSMUST00000134303	protein_coding	-1,319	0,00618	-1,139	0,00001
Slc51a	protein_coding	ENSMUST00000232418	retained_intron	-1,634	0,00169	-1,633	0,00008
Slc5a12	protein_coding	ENSMUST00000111026	protein_coding	-1,530	0,00087	-3,383	0,00001
Slc5a2	protein_coding	ENSMUST00000146735	retained_intron	-1,090	0,00082	-3,704	0,00092
Slc5a2	protein_coding	ENSMUST00000171335	retained_intron	-1,234	0,00081	-3,279	0,00003
Slc5a4a	protein_coding	ENSMUST00000020450	protein_coding	-2,482	0,00899	-3,757	0,00117
Slc5a8	protein_coding	ENSMUST00000020255	protein_coding	-1,332	0,00006	-2,551	0,00000
Slc6a19	protein_coding	ENSMUST00000132085	retained_intron	-1,327	0,00206	-1,530	0,00038
Slc6a19	protein_coding	ENSMUST00000022048	protein_coding	-1,394	0,00027	-1,879	0,00003
Slc6a20b	protein_coding	ENSMUST00000166800	nonsense_mediated_decay	-1,080	0,00081	-2,060	0,00000
Slc6a20b	protein_coding	ENSMUST00000171122	processed_transcript	-1,270	0,00011	-2,888	0,00002
Slc6a20b	protein_coding	ENSMUST00000026273	protein_coding	-1,352	0,00013	-2,232	0,00000
Slc7a13	protein_coding	ENSMUST00000035890	protein_coding	-1,328	0,00022	-4,836	0,00000
Slc7a7	protein_coding	ENSMUST00000195970	protein_coding	-1,019	0,00027	-2,901	0,00001
Slc9a8	protein_coding	ENSMUST00000148454	processed_transcript	-2,049	0,00014	-2,602	0,00001
Slco1a1	protein_coding	ENSMUST00000042119	protein_coding	-1,077	0,00507	-3,046	0,00013
Slco1a1	protein_coding	ENSMUST00000168119	protein_coding	-1,193	0,00082	-2,402	0,00030
Slco1a6	protein_coding	ENSMUST00000111827	protein_coding	-1,756	0,00011	-3,215	0,00000
Slit2	protein_coding	ENSMUST00000173107	protein_coding	-1,357	0,00139	-2,665	0,00097
Smarca2	protein_coding	ENSMUST00000176030	protein_coding	-1,583	0,00049	-1,843	0,00003
Smyd3	protein_coding	ENSMUST00000131684	processed_transcript	-4,997	0,00062	-1,272	0,00081

APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(continuação)

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Snx29	protein_coding	ENSMUST00000150993	protein_coding	-1,259	0,00404	-1,460	0,00000
Snx29	protein_coding	ENSMUST00000180792	protein_coding	-1,816	0,00012	-1,443	0,00000
Snx29	protein_coding	ENSMUST00000122168	protein_coding	-1,942	0,00115	-2,347	0,00000
Snx29	protein_coding	ENSMUST00000096273	protein_coding	-2,834	0,00432	-2,293	0,00007
Snx31	protein_coding	ENSMUST00000161202	protein_coding	-2,318	0,00389	-1,248	0,00588
Snx31	protein_coding	ENSMUST00000160936	processed_transcript	-2,994	0,00390	-2,706	0,00063
Sord	protein_coding	ENSMUST00000110551	protein_coding	-1,319	0,00006	-2,750	0,00000
Spp2	protein_coding	ENSMUST00000027518	protein_coding	-1,464	0,00012	-2,765	0,00001
Srd5a2	protein_coding	ENSMUST00000043458	protein_coding	-1,166	0,00128	-1,815	0,00257
St8sia1	protein_coding	ENSMUST00000032421	protein_coding	-1,468	0,00045	-2,575	0,00001
Stxbp5l	protein_coding	ENSMUST00000114787	protein_coding	-2,251	0,00996	-4,054	0,00700
Suc1g2	protein_coding	ENSMUST00000079847	protein_coding	-1,230	0,00006	-1,159	0,00000
Suc1g2	protein_coding	ENSMUST00000204224	protein_coding	-1,243	0,00007	-1,383	0,00000
Sugct	protein_coding	ENSMUST00000220514	protein_coding	-1,108	0,00047	-2,710	0,00001
Susd2	protein_coding	ENSMUST00000218896	retained_intron	-1,070	0,00716	-4,135	0,00837
Susd2	protein_coding	ENSMUST00000095541	protein_coding	-1,247	0,00006	-3,433	0,00000
Syn3	protein_coding	ENSMUST00000120638	protein_coding	-1,171	0,00285	-1,262	0,00009
Sypl2	protein_coding	ENSMUST00000141387	protein_coding	-1,065	0,00141	-1,927	0,00018
Tbc1d32	protein_coding	ENSMUST00000099739	protein_coding	-1,595	0,00009	-1,207	0,00015
Tbxas1	protein_coding	ENSMUST00000161360	retained_intron	-1,165	0,00215	-1,823	0,00002
Tef	protein_coding	ENSMUST00000164673	retained_intron	-1,035	0,00031	-1,602	0,00001
Tef	protein_coding	ENSMUST00000109553	protein_coding	-1,105	0,00017	-1,754	0,00000
Tenm4	protein_coding	ENSMUST00000107166	protein_coding	-2,128	0,00133	-2,824	0,00424
Them7	protein_coding	ENSMUST00000068813	protein_coding	-1,750	0,00021	-1,912	0,00001
Them7	protein_coding	ENSMUST00000131060	processed_transcript	-2,224	0,00153	-1,932	0,00002
Thrb	protein_coding	ENSMUST00000022304	protein_coding	-1,832	0,00064	-1,260	0,00006
Tln2	protein_coding	ENSMUST00000216799	retained_intron	-1,625	0,00342	-1,545	0,00003
Tln2	protein_coding	ENSMUST00000214859	protein_coding	-1,670	0,00014	-4,983	0,00005
Tln2	protein_coding	ENSMUST00000040025	protein_coding	-1,735	0,00018	-2,352	0,00000
Tmco3	protein_coding	ENSMUST00000210402	retained_intron	-1,081	0,00037	-1,679	0,00002
Tmco3	protein_coding	ENSMUST00000210600	protein_coding	-1,118	0,00801	-1,549	0,00013
Tmem178b	protein_coding	ENSMUST00000180588	retained_intron	-4,281	0,00178	-2,754	0,00034
Tmem25	protein_coding	ENSMUST00000132020	nonsense_mediated_decay	-1,309	0,00053	-2,554	0,00020
Tmem25	protein_coding	ENSMUST00000213203	nonsense_mediated_decay	-1,431	0,00149	-2,151	0,00003
Tmem26	protein_coding	ENSMUST00000080995	protein_coding	-1,662	0,00030	-1,831	0,00000
Tnfaip8	protein_coding	ENSMUST00000155403	retained_intron	-2,088	0,00218	-4,622	0,00001
Tns1	protein_coding	ENSMUST00000187281	protein_coding	-1,257	0,00391	-1,192	0,00001
Tpk1	protein_coding	ENSMUST00000067888	protein_coding	-2,941	0,00013	-2,237	0,00008
Tprkb	protein_coding	ENSMUST00000202588	retained_intron	-1,648	0,00023	-1,788	0,00004
Tprkb	protein_coding	ENSMUST00000202142	processed_transcript	-2,224	0,00285	-2,024	0,00114
Trabd2b	protein_coding	ENSMUST00000094894	protein_coding	-1,600	0,00025	-1,470	0,00000
Trappc9	protein_coding	ENSMUST00000023276	protein_coding	-2,476	0,00021	-1,434	0,00003
Trappc9	protein_coding	ENSMUST00000089770	protein_coding	-3,648	0,00130	-2,213	0,00002
Trim63	protein_coding	ENSMUST00000135576	processed_transcript	-1,221	0,00095	-3,661	0,00928
Triqk	protein_coding	ENSMUST00000183345	protein_coding	-1,109	0,00194	-1,090	0,00630
Ttc36	protein_coding	ENSMUST00000044694	protein_coding	-1,439	0,00017	-4,037	0,00000
Ttc36	protein_coding	ENSMUST00000216449	retained_intron	-1,471	0,00065	-4,333	0,00204

**APÊNDICE C – Transcritos subexpressos em comum ao Estudo 1 e Estudo 2
(final)**

Gene name	Gene type	Transcript	Transcript type	Estudo 1		Estudo 2	
				Fold change (log2)	Q-value	Fold change (log2)	Q-value
Ttc39c	protein_coding	ENSMUST00000234966	protein_coding	-1,080	0,00044	-1,174	0,00030
Ttc41	protein_coding	ENSMUST00000075632	protein_coding	-1,760	0,00506	-1,078	0,00060
Tyw3	protein_coding	ENSMUST00000052774	protein_coding	-1,489	0,00026	-2,024	0,00001
Ugt2b37	protein_coding	ENSMUST00000075858	protein_coding	-1,427	0,00026	-4,002	0,00000
Ugt2b38	protein_coding	ENSMUST00000072818	protein_coding	-1,633	0,00041	-3,923	0,00000
Ugt3a1	protein_coding	ENSMUST00000022861	protein_coding	-1,263	0,00016	-3,740	0,00000
Ugt3a2	protein_coding	ENSMUST00000072403	protein_coding	-1,156	0,00007	-3,022	0,00000
Unc5c	protein_coding	ENSMUST00000106236	protein_coding	-1,515	0,00120	-1,707	0,00001
Upb1	protein_coding	ENSMUST00000138058	retained_intron	-1,048	0,00042	-2,684	0,00034
Upb1	protein_coding	ENSMUST00000039925	protein_coding	-1,255	0,00007	-2,043	0,00000
Upb1	protein_coding	ENSMUST00000129946	retained_intron	-1,355	0,00206	-2,218	0,00014
Upp2	protein_coding	ENSMUST00000229923	protein_coding	-2,682	0,00411	-4,291	0,00386
Viil	protein_coding	ENSMUST00000051386	protein_coding	-1,282	0,00070	-1,172	0,00001
Vps8	protein_coding	ENSMUST00000118138	processed_transcript	-2,589	0,00021	-1,780	0,00533
Vps8	protein_coding	ENSMUST00000150765	retained_intron	-2,931	0,00090	-4,944	0,00011
Vps8	protein_coding	ENSMUST00000127903	processed_transcript	-5,630	0,00060	-1,573	0,00479
Vti1a	protein_coding	ENSMUST00000223690	protein_coding	-1,700	0,00021	-1,079	0,00034
Vti1a	protein_coding	ENSMUST00000095950	protein_coding	-1,772	0,00012	-1,000	0,00005
Vwa2	protein_coding	ENSMUST00000026068	protein_coding	-1,719	0,00025	-1,390	0,00000
Vwa8	protein_coding	ENSMUST00000040990	protein_coding	-1,693	0,00006	-1,118	0,00001
Vwce	protein_coding	ENSMUST00000055115	protein_coding	-1,814	0,00005	-1,901	0,00000
Wdpcp	protein_coding	ENSMUST00000020568	protein_coding	-2,376	0,00006	-1,482	0,00002
Wdr25	protein_coding	ENSMUST00000167816	protein_coding	-1,579	0,00673	-1,075	0,00026
Wfdc16	protein_coding	ENSMUST00000109336	protein_coding	-1,922	0,00053	-3,325	0,00004
Wwox	protein_coding	ENSMUST00000004756	protein_coding	-2,755	0,00008	-2,728	0,00000
Xylt1	protein_coding	ENSMUST00000032892	protein_coding	-2,954	0,00005	-2,096	0,00001

**APÊNDICE D – PROCESSOS BIOLÓGICOS ENRIQUECIDOS NOS
TRANSCRITOS CODIFICADORES DE PROTEÍNAS SUPEREXPRESSOS**

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0008150	biological_process	4375	240	219.53	1.3e-12
GO:0042771	intrinsic apoptotic signaling pathway in respo	19	9	0.95	3.9e-8
GO:0070059	intrinsic apoptotic signaling pathway in respons	27	8	1.35	8.6e-7
GO:0034644	cellular response to UV	33	10	1.66	1E-06
GO:0071480	cellular response to gamma radiation	15	7	0.75	1.6e-6
GO:0008630	intrinsic apoptotic signaling pathway in respo	50	18	2.51	8E-06
GO:0043525	positive regulation of neuron apoptotic process	20	7	1	1.6e-5
GO:1902042	negative regulation of extrinsic apoptotic sign	14	6	0.7	1.7e-5
GO:0002526	acute inflammatory response	22	6	1.1	1.9e-5
GO:0048147	negative regulation of fibroblast proliferation	10	5	0.5	3.6e-5
GO:0007093	mitotic cell cycle checkpoint	33	10	1.66	3.8e-5
GO:0034599	cellular response to oxidative stress	87	13	4.37	4E-05
GO:0035914	skeletal muscle cell differentiation	17	6	0.85	6.2e-5
GO:0043516	regulation of DNA damage response, signal trar	17	6	0.85	6.2e-5
GO:0043066	negative regulation of apoptotic process	287	37	14.4	6.5e-5
GO:0043065	positive regulation of apoptotic process	181	28	9.08	7.7e-5
GO:0072593	reactive oxygen species metabolic process	85	16	4.27	2.1e-4
GO:0036003	positive regulation of transcription fro...	14	5	0.7	2.5e-4
GO:0006974	cellular response to DNA damage stimulus	250	35	12.54	3.1e-4
GO:0051258	protein polymerization	90	5	4.52	3.5e-4
GO:2001022	positive regulation of response to DNA damage	33	6	1.66	3.6e-4
GO:0007267	cell-cell signaling	324	18	16.26	4.1e-4
GO:0045597	positive regulation of cell differentiation	293	29	14.7	5.4e-4
GO:0071222	cellular response to lipopolysaccharide	46	8	2.31	6.2e-4
GO:0072422	signal transduction involved in DNA damage cl	10	4	0.5	6.6e-4
GO:0019730	antimicrobial humoral response	10	4	0.5	6.6e-4
GO:0032459	regulation of protein oligomerization	10	4	0.5	6.6e-4
GO:0090184	positive regulation of kidney development	10	4	0.5	6.6e-4
GO:0071158	positive regulation of cell cycle arrest	10	4	0.5	6.6e-4
GO:0090200	positive regulation of release of cytoch...	10	4	0.5	6.6e-4
GO:0060561	apoptotic process involved in morphogenesis	11	4	0.55	0.00101
GO:0001570	vasculogenesis	28	6	1.4	0.00124
GO:0055114	oxidation-reduction process	264	21	13.25	0.00131
GO:2001235	positive regulation of apoptotic signaling path	61	11	3.06	0.00135
GO:0001775	cell activation	222	25	11.14	0.00136
GO:0033598	mammary gland epithelial cell proliferation	12	4	0.6	0.00146
GO:0051187	cofactor catabolic process	12	4	0.6	0.00146
GO:0051050	positive regulation of transport	282	22	14.15	0.00147
GO:0006749	glutathione metabolic process	20	5	1	0.00154
GO:0030308	negative regulation of cell growth	70	9	3.51	0.00171
GO:0046889	positive regulation of lipid biosynthetic process	21	5	1.05	0.00195
GO:0071347	cellular response to interleukin-1	21	5	1.05	0.00195
GO:0006949	syncytium formation	20	5	1	0.00196
GO:0046364	monosaccharide biosynthetic process	28	5	1.4	0.00197
GO:0097237	cellular response to toxic substance	45	7	2.26	0.00198
GO:0042304	regulation of fatty acid biosynthetic process	13	4	0.65	0.00203
GO:0097755	positive regulation of blood vessel diameter	13	4	0.65	0.00203
GO:0060713	labyrinthine layer morphogenesis	13	4	0.65	0.00203
GO:0034113	heterotypic cell-cell adhesion	13	4	0.65	0.00203
GO:0008285	negative regulation of cell population proliferat	175	25	8.78	0.0026

**APÊNDICE E – PROCESSOS BIOLÓGICOS ENRIQUECIDOS NOS
TRANSCRITOS PROCESSADOS ORIUNDOS DE GENES CODIFICADORES DE
PROTEÍNAS SUPEREXPRESSOS**

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0031110	regulation of microtubule polymerization or depolym	11	3	0.52	0.0055
GO:0001578	microtubule bundle formation	13	3	0.61	0.0091
GO:0051896	regulation of protein kinase B signaling	14	3	0.66	0.0113
GO:0007050	cell cycle arrest	14	3	0.66	0.0113
GO:0002831	regulation of response to biotic stimulus	16	3	0.75	0.0166
GO:0030516	regulation of axon extension	16	3	0.75	0.0166
GO:0030308	negative regulation of cell growth	18	3	0.85	23
GO:0007275	multicellular organism development	405	21	19.06	0.0283
GO:0048729	tissue morphogenesis	48	4	2.26	0.0289
GO:0014706	striated muscle tissue development	32	3	1.51	0.0345
GO:0048639	positive regulation of developmental growth	21	3	0.99	0.0348
GO:0099536	synaptic signaling	68	2	3.2	0.0361
GO:0006814	sodium ion transport	19	3	0.89	0.0364
GO:0010942	positive regulation of cell death	57	4	2.68	0.0368
GO:0030307	positive regulation of cell growth	22	3	1.04	0.0393
GO:0050848	regulation of calcium-mediated signaling	10	2	0.47	0.0455
GO:0071222	cellular response to lipopolysaccharide	10	2	0.47	0.0455
GO:0071236	cellular response to antibiotic	10	2	0.47	0.0455
GO:0061041	regulation of wound healing	10	2	0.47	0.0455
GO:0001909	leukocyte mediated cytotoxicity	10	2	0.47	0.0455
GO:0097191	extrinsic apoptotic signaling pathway	10	2	0.47	0.0455
GO:0015908	fatty acid transport	10	2	0.47	0.0455
GO:0051092	positive regulation of NF-kappaB transcription factor	10	2	0.47	0.0455
GO:1901880	negative regulation of protein depolymerization	10	2	0.47	0.0455
GO:0002708	positive regulation of lymphocyte mediated immunity	10	2	0.47	0.0455
GO:0031341	regulation of cell killing	10	2	0.47	0.0455
GO:0008150	biological_process	1395	53	65.64	0.0542
GO:0002822	regulation of adaptive immune response based on so	11	2	0.52	0.0544
GO:0000045	autophagosome assembly	11	2	0.52	0.0544
GO:0030335	positive regulation of cell migration	42	4	1.98	0.0566
GO:0055086	nucleobase-containing small molecule metabolic proc	55	5	2.59	0.0605
GO:0010212	response to ionizing radiation	12	2	0.56	0.0638
GO:0030336	negative regulation of cell migration	12	2	0.56	0.0638
GO:0014066	regulation of phosphatidylinositol 3-kinase signaling	12	2	0.56	0.0638
GO:0050772	positive regulation of axonogenesis	12	2	0.56	0.0638
GO:0007093	mitotic cell cycle checkpoint	12	2	0.56	0.0638
GO:0050709	negative regulation of protein secretion	12	2	0.56	0.0638
GO:0030001	metal ion transport	60	5	2.82	0.0664
GO:0006955	immune response	87	7	4.09	0.0676
GO:0050776	regulation of immune response	42	3	1.98	0.0683
GO:0050865	regulation of cell activation	30	2	1.41	69
GO:0098660	inorganic ion transmembrane transport	45	2	2.12	0.0697
GO:0042060	wound healing	29	4	1.36	0.0702
GO:0045892	negative regulation of transcription, DNA-templated	93	4	4.38	0.0702
GO:0072331	signal transduction by p53 class mediator	13	2	0.61	0.0737
GO:0030218	erythrocyte differentiation	13	2	0.61	0.0737
GO:0046651	lymphocyte proliferation	13	2	0.61	0.0737
GO:0046496	nicotinamide nucleotide metabolic process	14	2	0.66	0.0841
GO:0007369	gastrulation	14	2	0.66	0.0841
GO:0072593	reactive oxygen species metabolic process	14	2	0.66	0.0841

**APÊNDICE F – PROCESSOS BIOLÓGICOS ENRIQUECIDOS NOS
TRANSCRITOS CODIFICADORES DE PROTEÍNAS SUBEXPRESSOS**

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0008150	biological_process	5667	349	321.2	7.6e-24
GO:0055114	oxidation-reduction process	413	50	23.41	1.6e-10
GO:0042738	exogenous drug catabolic process	12	8	0.68	8.9e-9
GO:1901606	alpha-amino acid catabolic process	34	13	1.93	2.7e-8
GO:0006082	organic acid metabolic process	441	77	25	3.9e-7
GO:0017144	drug metabolic process	290	47	16.44	8E-07
GO:1901617	organic hydroxy compound biosynthetic p...	68	11	3.85	3.7e-6
GO:1901607	alpha-amino acid biosynthetic process	35	9	1.98	3.8e-6
GO:0008206	bile acid metabolic process	11	6	0.62	3.8e-6
GO:0006805	xenobiotic metabolic process	37	10	2.1	4.9e-6
GO:0042737	drug catabolic process	53	18	3	1.1e-5
GO:0055085	transmembrane transport	412	43	23.35	1.2e-5
GO:1901361	organic cyclic compound catabolic proces.	216	22	12.24	1.3e-5
GO:0015893	drug transport	63	16	3.57	3.3e-5
GO:0009072	aromatic amino acid family metabolic pro..	15	6	0.85	3.5e-5
GO:0006694	steroid biosynthetic process	56	9	3.17	4.2e-5
GO:0009083	branched-chain amino acid catabolic proc..	10	5	0.57	4.5e-5
GO:0072330	monocarboxylic acid biosynthetic process	113	18	6.4	4.8e-5
GO:0015711	organic anion transport	146	24	8.28	6.5e-5
GO:0008207	C21-steroid hormone metabolic process	11	5	0.62	8E-05
GO:0019439	aromatic compound catabolic process	201	18	11.39	1.5e-4
GO:0016999	antibiotic metabolic process	64	10	3.63	0,0002
GO:0042632	cholesterol homeostasis	39	8	2.21	3.6e-4
GO:0006637	acyl-CoA metabolic process	45	9	2.55	0,0005
GO:0014070	response to organic cyclic compound	234	19	13.26	6.5e-4
GO:0015872	dopamine transport	10	4	0.57	7.9e-4
GO:0035640	exploration behavior	10	4	0.57	7.9e-4
GO:0006633	fatty acid biosynthetic process	65	10	3.68	7.9e-4
GO:0072329	monocarboxylic acid catabolic process	65	10	3.68	0.00115
GO:0035725	sodium ion transmembrane transport	35	6	1.98	0.00118
GO:0098661	inorganic anion transmembrane transport	19	5	1.08	0.00147
GO:0030073	insulin secretion	67	10	3.8	0.00166
GO:0070328	triglyceride homeostasis	12	4	0.68	0.00173
GO:0048511	rhythmic process	98	10	5.55	0.00178
GO:0043648	dicarboxylic acid metabolic process	52	9	2.95	0.00185
GO:0008652	cellular amino acid biosynthetic process	37	11	2.1	0.00209
GO:0007586	digestion	31	5	1.76	0.00216
GO:0046700	heterocycle catabolic process	198	16	11.22	0.00231
GO:0007584	response to nutrient	40	7	2.27	0.00234
GO:0051187	cofactor catabolic process	21	5	1.19	0.00238
GO:0042398	cellular modified amino acid biosyntheti...	21	5	1.19	0.00238
GO:0035082	axoneme assembly	13	4	0.74	0.0024
GO:0009108	coenzyme biosynthetic process	93	15	5.27	0.0025
GO:0006814	sodium ion transport	78	12	4.42	0.00258
GO:1901362	organic cyclic compound biosynthetic pro.	1272	62	72.1	0.00281
GO:0006885	regulation of pH	34	4	1.93	0.00309
GO:0003341	cilium movement	15	4	0.85	0.00426
GO:0006767	water-soluble vitamin metabolic process	15	4	0.85	0.00426
GO:0046394	carboxylic acid biosynthetic process	162	32	9.18	0.00435
GO:1903510	mucopolysaccharide metabolic process	20	6	1.13	0.00467

**APÊNDICE G – PROCESSOS BIOLÓGICOS ENRIQUECIDOS NOS
TRANSCRITOS PROCESSADOS ORIUNDOS DE GENES CODIFICADORES DE
PROTEÍNAS SUPEREXPRESSOS**

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0055114	oxidation-reduction process	118	21	5.77	4.5e-10
GO:0043648	dicarboxylic acid metabolic process	21	8	1.03	1.9e-7
GO:1901606	alpha-amino acid catabolic process	19	7	0.93	1.5e-6
GO:0042737	drug catabolic process	18	6	0.88	1.8e-5
GO:0071466	cellular response to xenobiotic stimulus	14	5	0.68	6.5e-5
GO:0010038	response to metal ion	40	6	1.95	1.3e-4
GO:0055081	anion homeostasis	11	4	0.54	3.5e-4
GO:0009064	glutamine family amino acid metabolic process	11	4	0.54	3.5e-4
GO:0051384	response to glucocorticoid	12	4	0.59	5.1e-4
GO:1901568	fatty acid derivative metabolic process	13	4	0.64	7.2e-4
GO:1901361	organic cyclic compound catabolic process	80	9	3.91	8.6e-4
GO:0006635	fatty acid beta-oxidation	14	4	0.68	9.8e-4
GO:0006885	regulation of pH	15	4	0.73	0.00111
GO:0016999	antibiotic metabolic process	16	4	0.78	0.00169
GO:0008150	biological_process	2175	84	106.27	0.00233
GO:1901605	alpha-amino acid metabolic process	44	12	2.15	0.00293
GO:0006873	cellular ion homeostasis	75	5	3.66	0.00335
GO:1901617	organic hydroxy compound biosynthetic process	31	4	1.51	0.00385
GO:0042136	neurotransmitter biosynthetic process	11	3	0.54	0.00522
GO:0098771	inorganic ion homeostasis	91	8	4.45	0.00615
GO:0043434	response to peptide hormone	44	4	2.15	0.00653
GO:0019359	nicotinamide nucleotide biosynthetic process	16	3	0.78	0.00654
GO:0006090	pyruvate metabolic process	17	3	0.83	0.00655
GO:1901607	alpha-amino acid biosynthetic process	12	3	0.59	0.00679
GO:0016042	lipid catabolic process	45	10	2.2	0.00701
GO:0032787	monocarboxylic acid metabolic process	78	17	3.81	0.00781
GO:1901655	cellular response to ketone	13	3	0.64	0.00861
GO:0035725	sodium ion transmembrane transport	13	3	0.64	0.00861
GO:0009062	fatty acid catabolic process	19	6	0.93	0.00994
GO:0097327	response to antineoplastic agent	14	3	0.68	0.01069
GO:0071702	organic substance transport	356	13	17.39	0.01382
GO:0006637	acyl-CoA metabolic process	16	3	0.78	0.01566
GO:0006633	fatty acid biosynthetic process	17	3	0.83	0.01856
GO:0072330	monocarboxylic acid biosynthetic process	36	6	1.76	0.02024
GO:0009166	nucleotide catabolic process	19	3	0.93	0.02142
GO:0043604	amide biosynthetic process	119	5	5.81	0.02276
GO:0044282	small molecule catabolic process	63	15	3.08	0.02301
GO:0006694	steroid biosynthetic process	19	3	0.93	0.02518
GO:0046394	carboxylic acid biosynthetic process	56	12	2.74	0.02914
GO:0006520	cellular amino acid metabolic process	57	15	2.79	0.03308
GO:0046700	heterocycle catabolic process	75	7	3.66	0.03341
GO:2000060	positive reg. of ubiquitin-dependent protein catabolic	14	2	0.68	0.03397
GO:0042278	purine nucleoside metabolic process	14	2	0.68	0.03397
GO:0090066	regulation of anatomical structure size	70	3	3.42	0.03435
GO:0043068	positive regulation of programmed cell death	88	3	4.3	0.03461
GO:0042493	response to drug	96	10	4.69	0.03609
GO:0032774	RNA biosynthetic process	395	5	19.3	0.0388
GO:0019725	cellular homeostasis	121	10	5.91	0.04135
GO:0051453	regulation of intracellular pH	10	2	0.49	0.04352
GO:0050848	regulation of calcium-mediated signaling	10	2	0.49	0.04352

**APÊNDICE H – PROCESSOS BIOLÓGICOS ENRIQUECIDOS NOS GENES QUE
SOFRERAM TRANSCRIÇÃO *READTHROUGH* – DOGS SUPEREXPRESSOS**

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0008150	biological_process	6201	881	638.28	< 1e-30
GO:0000381	regulation of alternative mRNA splicing,...	36	15	3.71	9.5E-07
GO:0006886	intracellular protein transport	463	86	47.66	2.4E-05
GO:0008380	RNA splicing	201	48	20.69	6.4E-05
GO:1903828	negative regulation of cellular protein ...	61	18	6.28	0.00012
GO:0045292	mRNA cis splicing, via spliceosome	18	8	1.85	0.00021
GO:0048255	mRNA stabilization	18	8	1.85	0.00021
GO:0000045	autophagosome assembly	53	15	5.46	0.00026
GO:0061028	establishment of endothelial barrier	23	9	2.37	0.00027
GO:0016482	cytosolic transport	82	21	8.44	0.00033
GO:2000008	regulation of protein localization to ce...	21	8	2.16	0.00034
GO:0006397	mRNA processing	243	58	25.01	0.00035
GO:0061157	mRNA destabilization	15	7	1.54	0.00037
GO:0042742	defense response to bacterium	44	11	4.53	0.00041
GO:0051098	regulation of binding	179	38	18.42	0.00048
GO:0007034	vacuolar transport	73	17	7.51	0.0005
GO:0048010	vascular endothelial growth factor recep...	20	6	2.06	0.00051
GO:0060284	regulation of cell development	428	68	44.06	0.00055
GO:1904467	regulation of tumor necrosis factor secr...	12	6	1.24	0.00063
GO:0034067	protein localization to Golgi apparatus	21	8	2.16	0.00073
GO:2000300	regulation of synaptic vesicle exocytosi...	31	10	3.19	0.00075
GO:0008285	negative regulation of cell proliferatio...	256	43	26.35	0.00077
GO:0007032	endosome organization	32	10	3.29	0.00099
GO:0060048	cardiac muscle contraction	44	6	4.53	0.00102
GO:0032482	Rab protein signal transduction	27	9	2.78	0.00105
GO:0039694	viral RNA genome replication	13	6	1.34	0.00106
GO:0006376	mRNA splice site selection	18	7	1.85	0.00137
GO:0000122	negative regulation of transcription by ...	360	55	37.06	0.00164
GO:0031047	gene silencing by RNA	58	13	5.97	0.00167
GO:0061912	selective autophagy	28	9	2.88	0.00168
GO:0033962	cytoplasmic mRNA processing body assembl...	14	6	1.44	0.0017
GO:0007030	Golgi organization	74	17	7.62	0.00172
GO:0015031	protein transport	809	138	83.27	0.00177
GO:0043162	ubiquitin-dependent protein catabolic pr...	10	5	1.03	0.00185
GO:0032924	activin receptor signaling pathway	10	5	1.03	0.00185
GO:0060547	negative regulation of necrotic cell dea...	10	5	1.03	0.00185
GO:0006890	retrograde vesicle-mediated transport, G...	29	9	2.99	0.00186
GO:0032388	positive regulation of intracellular tra...	111	23	11.43	0.00191
GO:0050772	positive regulation of axonogenesis	47	11	4.84	0.00197
GO:0043525	positive regulation of neuron apoptotic ...	35	10	3.6	0.00212
GO:0051865	protein autoubiquitination	42	11	4.32	0.00276
GO:0030218	erythrocyte differentiation	61	11	6.28	0.00279
GO:0043900	regulation of multi-organism process	169	33	17.4	0.00287
GO:0045944	positive regulation of transcription by ...	455	70	46.83	0.00296
GO:0050663	cytokine secretion	78	16	8.03	0.00304
GO:0046824	positive regulation of nucleocytoplasmic...	30	9	3.09	0.00307
GO:1900027	regulation of ruffle assembly	11	5	1.13	0.00311
GO:0001701	in utero embryonic development	223	36	22.95	0.00312
GO:0043433	negative regulation of DNA-binding trans...	77	15	7.93	0.0033
GO:0051260	protein homooligomerization	152	27	15.65	0.00349

**APÊNDICE I – PROCESSOS BIOLÓGICOS ENRIQUECIDOS NOS GENES QUE
SOFRERAM TRANSCRIÇÃO *READTHROUGH* – DOGS SUBEXPRESSOS**

GO.ID	Term	Annotated	Significant	Expected	Fisher.weight01
GO:0008150	biological_process	5996	341	250.35	2.7E-20
GO:0055114	oxidation-reduction process	373	47	15.57	2.7E-09
GO:0042737	drug catabolic process	45	13	1.88	8E-09
GO:0008206	bile acid metabolic process	11	7	0.46	5.9E-08
GO:1901617	organic hydroxy compound biosynthetic pr...	77	13	3.22	2.7E-07
GO:0072330	monocarboxylic acid biosynthetic process	106	18	4.43	3.8E-06
GO:0006082	organic acid metabolic process	402	59	16.78	9.2E-06
GO:0005975	carbohydrate metabolic process	230	28	9.6	9.8E-06
GO:0006694	steroid biosynthetic process	60	13	2.51	1.3E-05
GO:0006090	pyruvate metabolic process	56	11	2.34	4E-05
GO:0006085	acetyl-CoA biosynthetic process	12	5	0.5	7.7E-05
GO:0006814	sodium ion transport	79	9	3.3	0.00014
GO:0032787	monocarboxylic acid metabolic process	225	39	9.39	0.00057
GO:0006637	acyl-CoA metabolic process	39	11	1.63	0.00063
GO:0006099	tricarboxylic acid cycle	18	5	0.75	0.00067
GO:0062013	positive regulation of small molecule me...	62	6	2.59	0.00078
GO:0010823	negative regulation of mitochondrion org...	29	6	1.21	0.00107
GO:0055085	transmembrane transport	440	34	18.37	0.00114
GO:0046496	nicotinamide nucleotide metabolic proces...	79	11	3.3	0.00126
GO:0006805	xenobiotic metabolic process	31	6	1.29	0.00154
GO:0008053	mitochondrial fusion	13	4	0.54	0.00158
GO:0008207	C21-steroid hormone metabolic process	13	4	0.54	0.00158
GO:0007586	digestion	28	5	1.17	0.00172
GO:0015908	fatty acid transport	31	5	1.29	0.00172
GO:0042493	response to drug	295	23	12.32	0.00176
GO:1990573	potassium ion import across plasma membr...	14	4	0.58	0.00214
GO:0090140	regulation of mitochondrial fission	14	4	0.58	0.00214
GO:0006631	fatty acid metabolic process	148	22	6.18	0.00281
GO:0033144	negative regulation of intracellular ste...	15	4	0.63	0.00283
GO:0035601	protein deacylation	60	6	2.51	0.00343
GO:0006635	fatty acid beta-oxidation	38	6	1.59	0.00451
GO:1901606	alpha-amino acid catabolic process	27	5	1.13	0.00458
GO:0007260	tyrosine phosphorylation of STAT protein	19	3	0.79	0.00506
GO:0050810	regulation of steroid biosynthetic proce...	28	5	1.17	0.00548
GO:0042407	cristae formation	10	3	0.42	0.00695
GO:0042159	lipoprotein catabolic process	10	3	0.42	0.00695
GO:0009311	oligosaccharide metabolic process	10	3	0.42	0.00695
GO:0007368	determination of left/right symmetry	49	6	2.05	0.00698
GO:0000186	activation of MAPKK activity	20	4	0.84	0.00851
GO:0006091	generation of precursor metabolites and ...	180	20	7.52	0.00853
GO:0006855	drug transmembrane transport	31	5	1.29	0.00855
GO:0009150	purine ribonucleotide metabolic process	192	26	8.02	0.00895
GO:1901361	organic cyclic compound catabolic proces...	202	16	8.43	0.00971
GO:0001676	long-chain fatty acid metabolic process	21	4	0.88	0.01017
GO:0007626	locomotory behavior	81	9	3.38	0.01135
GO:0015986	ATP synthesis coupled proton transport	12	3	0.5	0.01198
GO:0007176	regulation of epidermal growth factor-ac...	12	3	0.5	0.01198
GO:0045777	positive regulation of blood pressure	12	3	0.5	0.01198
GO:2000272	negative regulation of signaling recepto...	12	3	0.5	0.01198
GO:1900273	positive regulation of long-term synapti...	12	3	0.5	0.01198

APÊNDICE J – PUBLICAÇÃO RELACIONADA À TESE

Biomarkers and *in vitro* strategies for nephrotoxicity and renal disease assessment

Stellamaris Soares^{1,2}, Larissa C R Souza¹, Mark T Cronin³, Anna M Waaga-Gasser⁴, Marina F Grossi¹, Glória R Franco² and Carlos A Tagliati^{1*}

¹Departamento de Análises Clínicas e Toxicológicas, Faculdade de Farmácia, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

²Departamento de Bioquímica e Imunologia, ICB, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

³School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, England, UK

⁴Renal Division, Brigham and Women's Hospital, Boston, USA

Abstract

Acute kidney injury (AKI) is a global public health concern, impacting nearly 13.3 million patients and resulting in three million deaths per year. Chronic kidney disease (CKD) has increased by 135% since 1990, representing the pathology with the fastest growth rate worldwide. The annual costs of dialysis and kidney transplants range between US\$35,000 and US\$100,000 per patient. Despite its great impact, kidney disease has remained mostly asymptomatic for many years. AKI continues to be a major, unmet medical condition for which there are no pharmacological treatments available, while animal models are limited to provide direction for therapeutic translation into humans. Currently, serum creatinine is the standard biomarker to identify nephrotoxicity; however, it is a late stage biomarker. Hence, there is a pressing need to study *in vitro* biomarkers for the assessment of nephrotoxicity in order to develop new and safer drugs. Understanding of the mechanisms by which molecules produce nephrotoxicity is vital in order to both prevent adversity and treat kidney injury. In this review, we address new technologies and models that may be used to identify earlier biomarkers and pathways involved in nephrotoxicity, such as cell culture, omic approaches, bioinformatics platform, CRISPR/Cas9 genome-editing, *in silico*, organoids and 3D bioprinting, considering adverse outcome pathways (AOP).

Introduction

Kidney disease is a relatively silent disease and, in most cases, only becomes apparent to those suffering from it in the later stages of the illness. This can be partially attributed to the fact that most people are not aware of the early signs associated with kidney disease, coupled with the fact that these signs are not always recognized by healthcare providers [1]. Nonetheless, it is estimated that, annually, around three million deaths worldwide are caused by acute kidney injury (AKI). However, these numbers may be underestimated, given that half of the patients admitted to intensive care units develop AKI during hospitalization, and 27% of these die before being released from the hospital. It is also well-known that these figures constitute only a small proportion of the total number of reported cases [2].

AKI is an expected event that occurs relatively quickly, such as a few hours or days, after organ failure. The full process is still not clear. AKI is characterized by a reduction in the glomerular filtration rate (GFR), which leads to an increase in serum creatinine (sCr) or blood urea nitrogen (BUN). However, measuring BUN levels is not a reliable factor to predict AKI [3].

The frequency of AKI has significantly increased in recent decades and is now recognized as a significant cause of Chronic Kidney Disease (CKD) and kidney failure in general [4]. AKI is characterized as kidney disease that persists up to 90 days; this disease is considered chronic CKD when it surpasses 90 days [2,5].

Between 19% and 26% of the cases of AKI among hospitalized patients are caused by medication-induced adverse effects that injure the kidney [6]. AKI is still a major, unmet medical problem for which no efficient pharmacological treatments exist [7]. Therefore, an

understanding of the mechanisms by which molecules may cause harm (such as those leading to AKI) is vital in order to both prevent toxicity and treat kidney injury.

Acknowledging the need for action, an increasing number of global advocacy initiatives, such as World Kidney Day, the International Society of Nephrology 0by25, and the Lancet Kidney Campaign aim to raise public awareness regarding the consequences, costs, and impact of both CKD and AKI [2].

Acute kidney injury

AKI is a worldwide public health concern, impacting nearly 13.3 million patients per year [7]. AKI is associated with high morbidity and mortality - over 1.7 million deaths per year among adults and children worldwide [8]. AKI is disseminated throughout 72 countries and is higher in developing countries. The mortality rates of critically ill patients with AKI in hospitals ranges between 20% and 70% [9,10]. In the United States alone, half a million hospital admissions each year are complicated by AKI, resulting in total costs of approximately \$10 billion [11]. Worldwide, only half of the individuals who need renal replacement therapy can be treated, due to difficulties related to access to care or lack of diagnosis. In fact, estimates suggest that 2.5 to 5

*Correspondence to: Carlos Alberto Tagliati, ToxLab, Departamento de Análises Clínicas e Toxicológicas, Faculdade de Farmácia, Universidade Federal de Minas Gerais, Brazil, Tel: 553134096891; E-mail: carlostagliati@ufmg.br

Key words: acute kidney injury (AKI), adverse outcome pathways (AOP), chronic kidney disease (CKD), early diagnosis, nephrotoxicity models, new technologies

Received: March 02, 2020; **Accepted:** April 17, 2020; **Published:** April 20, 2020

million patients are not properly treated [12]. In addition, the costs of dialysis and kidney transplants range between \$35,000 and \$100,000 per patient annually.

Prescription drugs and other traditional low-level medication is the major cause of AKI, resulting in nearly 25% of all patients who use at least one nephrotoxic drug. The other 75% of the cases stem from other etiologies. However, the main cause of mortality and morbidity by AKI can be explained by drug-induced nephrotoxicity in over 20% of hospital-acquired AKIs [13-16]. Despite this high rate, data show that only 7% of new drug candidates fail in preclinical trials because of nephrotoxicity [17], whereas the incidence of patients in intensive care units developing AKI is about 30-50% [18]. This discrepancy may help to explain the underestimation of nephrotoxicity in preclinical trials [19]. Therefore, AKI is still a major medical condition for which no pharmacological treatment has been developed [20-22].

The incidence of AKI has greatly increased over the past two decades, and it is now recognized as an important cause of CKD and kidney failure [4,23]. However, the causal pathway leading to CKD is complex and will require a broad multidisciplinary response [24]. Fortunately, the kidney carries an extremely regenerative potential and, in some cases, can be completely recovered '*restitutio at integrum*' after an AKI. Nevertheless, there is strong evidence suggesting that repair can be incomplete, and an episode of AKI may lead to CKD. The appearance of tubular response after AKI will shed light on the connection between incomplete repair and future risk of CKD [3]. Thus, the 'AKI to CKD transition' might involve incomplete renal tubule repair after AKI, eventually triggering interstitial renal fibrosis, making it critical to investigate the cells responsible for normal repair. Hence, by developing new targeted therapies to reduce injury and promote healing in AKI, it will be possible to determine which cells should be targeted.

In clinical practice, the early detection of AKI has had no impact on clinical decisions. In addition, therapeutic possibilities for AKI are very limited, and it is unclear if there is any benefit in either early or late onset of renal replacement therapy. As a result, the early or late detection of AKI, in most cases, does not impact the clinical course of the patient. To some extent, this may explain why most biomarkers have not been included in the panel of classic laboratory parameters used in the clinical setting. Hence, advances are needed in this area in order to understand the gap related to renal disease.

Chronic kidney disease

According to the "Kidney Disease: Improving Global Outcomes (KDIGO)" guidelines, CKD is characterized by abnormalities of kidney structure or function that are present for more than 3 months [5].

There are few markers, such as neutrophil gelatinase-associated lipocalin (NGAL), kidney injury molecule-1 (KIM-1), and nephrin (NPNT), as well as the validated markers of proteinuria and microalbuminuria, that can detect the progression of AKI into CKD, which makes it an important area for investigation [25,26]. Among those who survive the long-term outcomes of AKI, patients with CKD can progress into the end-stages of renal disease (ESRD) or into the aggravation of pre-existing CKD, which can quickly lead to ESRD [12,23,27].

Thus, AKI, which was previously thought to have no major consequences for recovered patients, may lead to poor life quality as well as long-term financial costs [28], which increases the need for public awareness regarding the early signs of the disease and investment

in further research [26]. CKD represent an important risk factor for cardiovascular and cerebrovascular diseases and progress toward ESRD [29,30], which can only be treated by renal replacement therapies such as hemodialysis, peritoneal dialysis, or transplants [30,31]. Kidney failure is the most severe form of CKD and is fatal if not treated by dialysis or kidney transplant. Because of the uncertainty of what causes CKD, many patients do not receive the correct treatment [4,32].

The resources required to treat CKD-associated complications (including kidney failure) impose a huge burden in such countries as Malaysia, Tunisia, Turkey, Chile, Mexico, and Uruguay [33]. In fact, it has been shown that the burden of CKD is on the rise, as indicated by the increase in both deaths and ESRD [4,23,34-36].

Nevertheless, CKD is not on the priority list of non-communicable diseases, and only a few countries have explicit policies or public programs focused on preventing and controlling CKD. This unfortunate scenario is of concern because there is a gap between recent advances in treatment and the increase in mortality worldwide, thus creating barriers that limit the progress of basic research that would ultimately lead to new treatments (translational research type 1) [37].

CKD represents the pathology with the fastest growing incidence worldwide [38]. Its incidence has virtually duplicated between 1990 to 2010 [39]. The global prevalence of CKD is estimated at around 10% to 15%, with similar estimates in men and women, and is more frequent in high-income countries. The impaired renal function is caused by diabetes mellitus in 30-40% of the cases, hypertension in 20%, inflammation of the glomeruli (glomerulonephritis) in 13%, interstitial nephritis in 10%, and polycystic kidney disease in 6%. Regardless of the cause, the common final outcome of almost all progressive CKDs is renal fibrosis [40,41].

The actual incidence of renal disease seems to be underestimated because many deaths attributed to diabetes have actually been caused by renal disease [42]. Moreover, hypertension and obesity are among the growing non-communicable diseases and represent important risk factors for CKD [23]. Thus, identifying CKD has become imperative not only to prevent the progression of the disease, but also to reduce the risk of cardiovascular morbidity and mortality. Hence, new biomarkers are paramount for early identification of the disease.

Biomarkers for nephrotoxicity

Currently, AKI diagnosis is based on serum creatinine's rise and/or fall in urine production. These are not renal-specific markers and have major limitations. Alternatively, there are two sets of biomarkers available for clinical use: NGAL and the combination of urinary insulin-like growth factor-binding protein 7 (IGFBP-7) and tissue inhibitor of metalloproteinases 2 (TIMP-2), two markers of cell-cycle arrest. These biomarkers seem to be specific and sensitive enough to be used together with additional markers for AKI to better stratify renal injury [43].

In 2008, the Food and Drug Administration (FDA) designated seven biomarkers of nephrotoxicity to test in animals and, on a case-by-case basis, to be used in humans. These included urinary KIM-1, β_2 -microglobulin, cystatin C, clusterin, trefoil factor 3, albumin, and total protein. These markers, among others (eg, urinary NGAL, urinary IL-18, and L-FABP - liver fatty acid binding protein), have been evaluated in many conditions but have rendered inconclusive results as a predictive model. In most cases, the data obtained with biomarkers tested in animals are compared to changes in sCr concentrations in humans, but results are questionable. In 2014, the NephroCheck

test, which helps to determine if critically ill patients are at risk of developing moderate-to-severe AKI following 12-hour testing, was approved by the FDA. However, it is unknown at which time-point the measurement should be performed to ideally predict AKI [26,44,45]. Thus, there are few markers for progression of AKI to CKD. Moreover, this strategy of using biomarkers also serves to determine a specific point in the course of AKI when the predictive value of the biomarker should be tested [11].

AKI is an extremely heterogeneous and complex condition. Even though there is an effort to identify AKI biomarkers that could be used for risk assessment, diagnosis, severity, and/or outcome of AKI with high predictive power, an ideal, universal AKI marker is unlikely to be discovered. Thus, considering the heterogeneous nature of AKI, a complex multimodal approach, including a detailed risk assessment and the implementation of new biomarkers, is advisable to prevent and manage AKI [46].

For this reason, several biomarkers that comprise the different aspects of AKI are required. Sensitive detection of an acute reduction in kidney function and injury in various renal structures, as well as the evaluation of the degree of renal injury, are the types of information provided by marker panels. Ideally, these markers should also be mediators of other disease pathways in AKI. Even better, these markers are the key to understanding the pathological mechanisms, to indicating the etiology of AKI, to providing references for future therapies, as well as to following therapy progress. Early biomarkers to estimate renal tissue damage is an advance that replaces the old low-sensitive methods for diagnosis by using better AKI criteria [7,47].

AKI biomarkers may be used to reveal higher sCr levels after ischemia reperfusion injury (IRI) and can also be used as an early indication of the progression from AKI to CKD. This may also indicate that the process used was not as effective as intended. For instance, fibrosis is still found in patients, despite the reacquisition of normal biochemical parameters, such as plasma creatinine removal; progression to CKD is still the most frequent result [48].

KIM-1 is a type I cell membrane glycoprotein with extracellular immunoglobulin and mucin domains [49-51]. The mRNA and protein for KIM-1 are expressed at very low levels in the normal rodent kidney; however, expressions increase dramatically after an injury in the proximal tubule epithelial cells in postischemic rodent kidney and in humans during ischemic acute renal ups and downs of using KIM-1, and according to Medić, *et al.* [52], larger trials are still necessary before KIM-1 can be tested in broader clinical trials. Although KIM-1 is a sensitive and specific marker for kidney tubular injury, it is difficult to measure it in acute settings [53].

By contrast, Ichimura, *et al.* [54] demonstrated that human KIM-1 exhibits homology to HAVCR1, and studies using gentamicin [55,56], amphotericin B [57], and cisplatin (no publishing data) have shown that the gene HAVCR1 (KIM1) might be used as *in vitro* biomarkers of nephrotoxicity.

To find new drugs for AKI and prevent its transition into CKD, it is crucial to better understand how tubular injury and repair work. Once this knowledge is acquired, it would be possible to use translational biomarkers to assess drug targets and their pathways. Furthermore, the expected biological effect of the drug can be revealed by specific biomarkers, i.e., whether the drug has reached the pathophysiological pathway for which it was developed and what is the expected clinical outcome [3,47,48].

Preclinical studies for nephrotoxicity

Animal model

Animal models have been widely used to understand the pathogenesis and underlying mechanisms of renal disease. As shown in the Figure. 1, mice and rats are the most commonly used to study nephrotoxicity and therapeutic targets, as well as to identify new biomarkers, mainly because these are easily bred and are relatively inexpensive to house and maintain [58]. Nonetheless, few advances have been achieved in understanding nephrotoxicity mechanisms or in identifying either a new biomarker or a panel of biomarkers. According to our search, monkeys are also used; however, regardless of the size and breed of the animals, results did not show a good predictability to identify adverse effect in humans.

In fact, the failure of existing animal models to accurately predict nephrotoxicity has been a major barrier to the development of safer drugs. Extensively studied for several decades, animal models of human AKI have, to date, yielded no specific therapy that benefits the human disease, either in preventing its occurrence, ameliorating its severity, hastening its recovery, or retarding the risk of ensuing CKD. Such inability to translate these models' uncovered salutary strategies into therapies for human AKI has led to a questioning of the value of these models: specifically, their resemblance and fidelity to human AKI. Overlapping this kidney-centered discussion are broader debates in relevant biomedical communities, national as well as international, regarding a worrisome, relative lack of congruence of conclusions in preclinical studies and the slow pace of translating findings obtained in animal models into new therapies for human disease [59].

Ischemia-reperfusion is one of the main causes of AKI; by definition, human AKI implies decreased GFR. Hence, conventional AKI models are mainly the result of kidney ischemia-reperfusion, exposure to endogenous or exogenous toxins and urinary tract obstruction [60]. However, this may not be the case for animal models if only one kidney is injured. Hence, despite the abundant literature on the pathogenesis of ischemia-reperfusion, no novel therapy has been able to be applied in clinical scenarios [61].

Exogenous drugs or poisons, as well as endogenous toxins, are widely used to stimulate AKI for their side or poisoning effects. Among these models, 6-20 mg/kg cisplatin can result in acute tubular injury within 72 h, whereas the administration of 40-200 mg/kg gentamicin in rats for 4-10 days can induce acute renal failure [62].

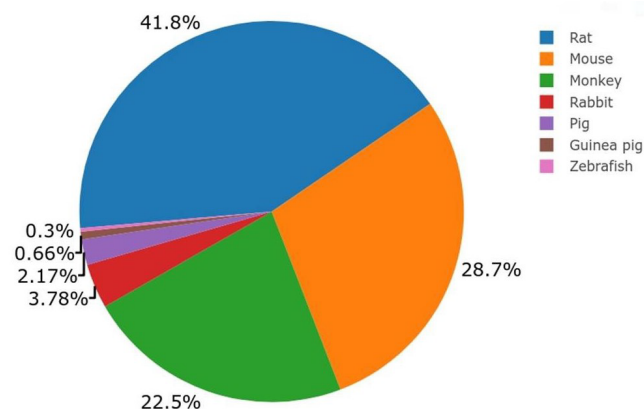


Figure 1. Proportion of animals used in *in vivo* studies in renal toxicity

Chemotherapy comprised three cycles of high dose (80-100 mg/m²) cisplatin administered on days 1, 22, and 43. Cisplatin is a drug widely used for chemotherapy (80-100 mg/m² – IV administered on days 1, 22, and 43) to effectively treat various cancers and gentamicin is an aminoglycoside antibiotic (6-12 g/day, IV administered, for 2-6 weeks) commonly used to prevent gram-negative bacterial infection. However, both drugs have limitations in clinical practice due to nephrotoxicity. For many years, nephropathy caused by cisplatin and gentamicin have been widely studied, but it is still unclear, although different models using different animals have been used and many hypotheses have been suggested [63-66]. Nevertheless, these hypotheses do not fully reproduce human clinical diseases. Hence, we do believe new *in vitro* strategies are necessary to study the molecular mechanism involved in nephropathy so as to reach proper clinical diagnoses and implement effective therapeutic interventions.

Actually, some AKI models in mice have been used with both advantages and disadvantages (Table. 1). Even so, models tend to have a high clinical relevance, for example, classical models contain a high knowledge background and are technically simple and reproducible. The types of models are incomplete and many details, such as model techniques and modeling time, are not mentioned. In addition, the majority of these have a variable response between models, expected acute renal necrosis is not always achieved, AKI is not produced clinically and pathologically, surgery is required, reproducible outcome is dependent on accurate Ischemia/Reperfusion time, and the renal function can be compensated by a non-ligated kidney [62].

Aristolochic acid (AA) and high dose folic acid (FA) are also frequently used to study AKI-CKD transition, with AKI models developed by warfarin and glycerol [62]. Nephrotoxicity caused by AA was first reported in Belgium in a group of patients who ingested weight loss pills containing extracts of Chinese herbal powdered roots [67]. The findings in humans are supported by laboratory animal studies showing that oral exposure to AA caused tumors in the urinary tracts of various animal species. Although this model showed good results it did not work for other drugs. In 2001, the FDA advised consumers to stop using any products that may contain AA, after seeing an increase in nephropathy, or kidney disease, among users [68].

Fink, *et al.* [69] demonstrated that single intravenous doses of FA in rats causes direct nephrotoxic effects; hence, it has been used to study AKI-CKD transition. Some studies have used intraperitoneal injection [60,70,71]. However, FA is the synthesized form of folate present in fortified foods and supplements, such as grains, and has a higher bioavailability than naturally occurring folate [72].

Thus, AA and FA are used at low doses for oral administrations in human and at high dosages and intravenous and peritoneal doses when studying nephrotoxicity in animals. Hence, AA and FA do not represent good models, since the profiles for causing nephrotoxicity are completely different.

Warfarin (WRN) therapy can result in AKI by causing glomerular hemorrhage and renal tubular obstruction by red blood cell (RBC) casts [73]. Analysis of more than 15,000 warfarin-treated patients showed that WRN affects approximately 33% of CKD patients and 16% of non-CKD patients who experienced an international normalized ratio >3.0. This study also identified that mortality rates in patients with WRN was significantly higher than in patients without WRN [74]. However, excessive anticoagulation in control animals was not associated with changes in creatinine levels, and renal morphology was normal. Therefore, Ware, *et al.* [75] developed a new method to study nephrotoxicity using an excessive amount of warfarin combined with a 5/6-nephrectomy model. Although this method could cause functional and morphological damage of the remnant kidney, it does serve to demonstrate the effect. Nevertheless, it does not guarantee the same molecular mechanism responsible for nephrotoxicity, which is the main element to be discovered.

AKI is also a common complication of rhabdomyolysis and accounts for the high mortality [62,76,77]. To reproduce the typical symptoms observed in humans, rats or mice are deprived of water for 24 h, after which a 8-10 mL/kg dose of 50% glycerol is administered in the hindlimb muscle [78,79]. We do not believe glycerol could be a good model, given that the model is focused on the symptoms, which is not the main point of our study. In addition, glycerol is used in human food at low doses. Thus, it does not represent a good model to study nephrotoxicity.

Table 1. Advantages and limitations of conventional AKI animals models

Model	Renal target	Design	Advantages	Limitations
Ischemia-reperfusion	Proximal tubule and endothelium	Clamping time: 30-45 min Reperfusion time: 24-48 hours	Pathology similar to human disease	Requires surgery
Drug induced	Cisplatin	Single IP injection (6-20 mg/kg) 72 hours	Pathology, timing, and drug doses similar to human disease	No clinical correlate
	Gentamicin	Serial IP injections (40-200 mg/kg/day) 4-10 days	Induces rapidly progressive CKD	No fully reproduce human clinical diseases
	Aristolochic acid	Serial IP injections (5 mg/kg/day) 5 days	Induces rapidly progressive CKD	No clinical correlate
	Folic acid	Single IP injection (250 mg/kg) 24-48 hours	Pathology findings similar to human AKI from other causes	No clinical correlate
	Glycerol	Single IM injection (8 mL/kg of glycerol 50%) 24-48 hours	Clinically relevant, resembling human rhabdomyolysis	Severe and focuses in the symptoms, not the mechanism
	Warfarin	Proximal tubule and glomerulus	5/6 nephrectomy for 3 weeks and 8 days on warfarin	Useful to study AKI caused by anticoagulants

Abbreviations: IP: Intraperitoneal; IM: Intramuscular.

Mechanisms of disease generation and progression in AKI and CKD have yet to be fully understood [80,81]. Although several clinical studies have investigated early stage predictive biomarkers of kidney disease, few have been applied in clinical practice [82]. CKD models can include mainly renal mass reduction, diabetic nephropathy, hypertension glomerular injury, IgA nephropathy, polycystic kidney disease, and chronic tubulointerstitial nephritis (Table. 2).

Renal mass reduction (remnant kidney model) is one of the most widely used models to study CKD. Subtotal (5/6) nephrectomy has been a mainstay of studies of progressive CKD. The right kidney is removed and the upper and lower poles (2/3 of the left kidney) are resected after ligation of the left renal artery. After surgery, activation of the renin-angiotensin system (RAS) can cause glomerular hypertension/hyperfiltration [61,62,83,84].

Glomerular injury may be acquired or hereditary. Acquired diabetic nephropathy (DN) is the leading cause of ESRD. There are many rodent models for diabetic nephropathy; however, none of them perfectly mimics the human disease [85]. The Animal Models of Diabetic Complications Consortium (AMDCC) defines the ideal rodent model of human DN and complications [86,87]. Several models of gene

targeting have reproduced mutations to generate murine models of hereditary podocyte-based nephropathies, including focal segmental glomerulosclerosis and Alport syndrome [61]. Hypertension-induced renal injury is usually investigated by using spontaneous hypertension. Additionally, uninephrectomy is required to promote significant kidney injury in the form of albuminuria without reduced GFR [88]. Chronic exposure to pressure-increasing doses of angiotensin II in rats and mice caused renal dysfunction [89].

Focal segmental glomerulosclerosis (FSGS) is a common primary glomerular injury, characterized mainly by podocyte loss, glomerular fibrosis, and marked proteinuria [90]. Although there is currently no primary FSGS model available, several secondary FSGS models have been established. Non-genetic models mimic acute glomerular injury but not the spontaneous slowly progressive FSGS observed in the clinic. Experimental podocyte depletion, using adriamycin or puromycin aminonucleoside, in turn leading to CKD, emphasizes the key role of podocytes in the preservation of the renal function [61,91].

IgA nephropathy (IgAN) is the most common human glomerulonephritis, but no specific therapy has been discovered. However, none of these models enabled the identification of the factors

Table 2. Advantages and limitations of experimental CKD animal models

Pathology	Model	Transition AKI-CKD	Advantages	Limitations	Comments
Renal mass reduction	5/6 nephrectomy (rats)	Depends on the mice strains	Equivalent to humans	Two surgery requires. The second one having a high mortality rate, especially in mice, and requiring technical expertise; mice strains are differentially responsive to renal mass reduction with respect to CKD progression; it is irreversible and the amount of kidney left to study is small and may be distorted by surgical procedure	No clinical translation of successful
Diabetic Nephropathy	Streptozotocin mice/rats; NOD mice BB-DP rat; ob/ob mice db/db mice; STZ-eNOS ^{-/-} ; db/db-eNOS/mice	No	Gene modified; commercially available; available on multiple strains; useful to study mechanisms of diabetic nephropathy. Reproduce human physiopathology	None of the many animal models perfectly replicate the human disease; expensive; some strains are infertile; cause modest albuminuria without decreased GFR	No clinical translation of successful
Hypertensive Nephropathy	SHR rats+UNX; angiotensin II infusion models	Depends on the dose/time	Highly relevant to hypertension nephropathy; useful to study AngII effect over kidney	SHR rats are more resistant to streptozotocin-induced diabetes; uninephrectomy is required to promote significant kidney injury; no progressive GFR loss; higher dose or long exposure time periods increase serum creatinine levels; high cost; slow progression	No clinical translation of successful; long exposure time periods stress animals
Acquired Glomerular Injury	Focal segmental Glomerulosclerosis: adriamycin (rat, mice) or puromycin (rat)	Depends on the strains	Mimic acute glomerular injury	Usually no adequate reproduction of slowly progressive human disease	No clinical translation of successful NOD
IgA Nephropathy	ddY mouse, HIGA mice; Uteroblobin-deficient mice		Reproduces human pathology	Mild disease development usually without progression towards end-stage renal disease; no progressive GFR decrease	No clinical translation of successful
Polycystic Disease	Genetically engineered mouse model	Yes, but no specific to renal disease	Useful to study PKD mechanisms and to find therapeutic targets; new insights into the molecular mechanisms of cystogenesis and associated progression of CKD; resulted in regulatory approval for Tolvaptan in Japan for human ADPKD	Mouse models usually lack the variety of human phenotypes	ARPKD mouse models usually lack the variety of human phenotypes
Chronic Tubulointerstitial Nephritis	Simple, dose-dependent decreased GFR; Possibly reversible (rats)	Model no specific for human	Simple, dose-dependent decreased GFR; possibly reversible	Adenine-fed male rats had a more severe decline in kidney; models CKD extrarenal complications; function and molecular changes	Adenine intoxication is not an issue in human
Hereditary Glomerular Injury	Pkd1 or Pkd2 gene engineered; Col4a43 gene knockout mouse; Alport syndrome	Depends on the strains	Reproduce features of human disease, including progressive GFR loss; develop proteinuria and renal failure	Difficulty to translation of genetic findings into improved patient care; high costs	Resulted in clinical recommendations
Stone Disease	Natural development (dogs and cats)	No information	Without external factors to cause disease	Few studies	Long-term

Abbreviations: GFR: glomerular filtration rate; SHR: Spontaneously hypertensive rat; BB-DP: BioBreeding-diabetes prone; NOD: nonobese diabetic; ob/ob: obese; db/db mouse: type 2 diabetes; DBA/2J: model for congenital experimental glaucoma; STZ: streptozotocin; eNOS: Endothelial nitric oxide synthase; UNX: Uninephrectomy; ddY mouse: outbreak; HIGA: Nephropathy with Hyperserum IgA; PKD: polycystic kidney disease; COL: collagen genes

that control the transition between disease onset and progression to ESRD. Thus, the underlying mechanism of IgAN is still not fully understood [61,62].

Polycystic kidney disease (PKD) includes a group of human monogenic disorders inherited in an autosomal dominant (ADPKD) or recessive (ARPKD) fashion. Genetically engineered mouse models, in which disease-causing genes were targeted, now mimic human ADPKD, ARPKD, and atypical PKD. Those models have already provided new insights into the molecular mechanisms of cystogenesis and the associated progression of CKD. However, ARPKD mouse models usually lack the variety of human phenotypes [61].

Primary chronic tubulointerstitial nephritis is associated with an immune-mediated infiltration of the kidney interstitium by inflammatory cells, which may progress to fibrosis. This is characterized by tubular atrophy, interstitial fibrosis, and interstitial inflammation in the absence of significant glomerular disease. Most animal models of chronic tubulointerstitial nephritis are secondary to ischemia, proteinuria, AKI, or other primary kidney diseases [61,92]. Thus, the processes involved are different and do not demonstrate similar mechanisms; hence, it does not represent a good model.

In addition to the animals already mentioned dogs and cats also have been suggested to study the calcium oxalate stone formation, considering that a naturally occurring animal model could be more useful. This strategy could represent a good model; however, the pathologic and etiologic mechanism surrounding calcium oxalate nephrolithiasis is incomplete [93].

Thus, despite the valuable new insights into kidney disease gained from existing models, many do not fully reproduce human clinical diseases. In the AKI, models are not stable and reproducible, while in the CKD, the disease is quite complex [62]. However, in addition to each model's specific limitations, all animal-model research into human diseases is ultimately restricted by the need to translate findings across species. This calls for the wider use of human-based models to complement and reduce the use of experimental *in vivo* research [94].

In this scenario, we do believe that new strategies and models without animals are required to better mimic processes involved in AKI and CKD. These models must mainly address molecular mechanisms, since the mere understanding of them may help to develop therapeutic interventions and clinical diagnoses, as well as to identify new biomarkers for nephrotoxicity. Hence, we also believe that new strategies are paramount for advances in this area. Moreover, alternative methods, such as those addressed in this review, represent important tools that can shift the paradigm of nephropathy and the way new, more effective, and safer molecules can be developed.

***In vitro* model**

In vitro toxicity studies, whether cell or tissue, have been used to study the mechanisms of action and toxic effects of drugs in order to assess efficacy and safety in humans regarding exposure to xenobiotics, and may also help to classify them according to their toxic capacity. The knowledge of the mechanism of toxic action involving a xenobiotic is of paramount importance, since it provides significant aid in the evaluation of predictive values, in addition to improving the selection of new drugs to treat specific pathologies [95-97].

The number of *in vivo* studies has grown proportionally to *in vitro* studies (Figure. 2) according to our search. Despite this, such models have demonstrated limitations in generating knowledge of nephrotoxic processes in humans. By contrast, there is an increasing

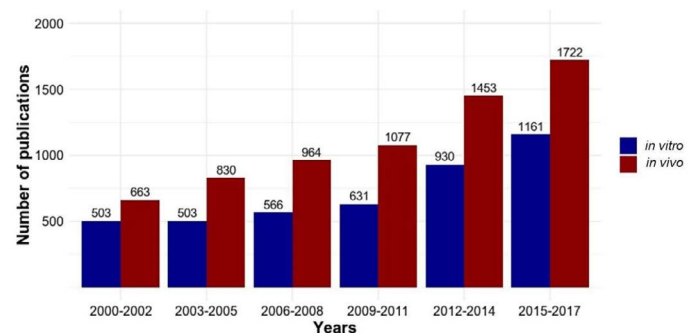


Figure 2. Number of *in vivo* and *in vitro* studies in renal toxicity

demand for new methods that refine, reduce, and replace animal use. Cell culture techniques are extremely relevant tools for *in vitro* studies of nephrotoxicity. The number of studies with animal models and *in vitro* renal toxicity has grown in the last 20 years. Despite this increase of both strategies, no reported results have provided a better understanding of nephrotoxicity mechanisms.

It is difficult to study specific epithelial cell types because the renal cortex carries a heterogeneous population of renal tubule segments. Fortunately, improvements in the methods used to grow homogeneous cultures of kidney cells allow one to access proper cells over different portions of the nephron. Other possibilities are the 3D printing of scaffolds to test the efficacy of new drugs, which will be different in scale and throughput from the assays used in the initial toxicological screening of chemicals [98].

Regardless of these advances, the kidneys produce very complex structures with blood filtration and urine re-absorption units needed for their function and homeostasis. It will also be highly important to consider the quality control of kidney organoids for disease modeling, added to the potential sources for developing kidney regenerative therapies. Conversely, there is currently no consensus regarding the methods that should be used for the evaluation of the quality of kidney organoids, mainly in terms of their functional characterization [99].

Recently, the mouse ENCODE Consortium reported that, while there are many similarities between human and mouse genomes, there are important differences as well. Consequently, animals are not good models for nephrotoxic effects in humans. Many discrepancies in DNA and differences in gene expression patterns have been found, potentially restricting the usefulness of some disease models based on mice [100]. More than 160 inherited genetic kidney diseases are now known.

Alternatively, by establishing human induced pluripotent stem cells (hiPSCs) from patients with a given genetic disease, such as degenerative disorders and cancer, it might be possible to study customized disease mechanisms and to perform drug screening *in vitro* instead of using animal models. Most importantly, updated advances in genome editing have provided new approaches to modeling genetic kidney diseases by using human pluripotent stem cells (hPSCs) *in vitro* [101-104].

In addition, the market potential for NATs (Non-Animal Tests) is enormous. Approximately \$100 billion was spent on R&D in 2014 by the top 25 pharmaceutical companies (based on recorded global sales for 2014) and a recent survey of the members of Pharmaceutical Research and Manufacturers of America estimated that more than 20% of total R&D expenditure was on preclinical research [98,105].

In the long term, there is the potential to use NAMs (Non-Animal Models) to support the development of personalized medicines through the use of human-based approaches that will enable the identification of the possibility for side effects caused by drugs, or variations of efficacy in the population. The selection of patients on the basis of their predicted response allows a drug to continue through its development into clinical use in circumstances where its progress would otherwise have been suspended. This could bring enormous benefits to both patients and companies [98].

In addition, there is an increasing demand for new methods which refine, reduce, and replace animal use. Cell culture techniques are extremely relevant tools for *in vitro* studies of nephrotoxicity. The number of studies with animal models and *in vitro* renal toxicity has grown in the last 20 years. And despite this increase in both strategies, it has not resulted in a better understanding of nephrotoxicity mechanisms (Figure. 2).

New technologies

While *in vitro* models have helped us to understand nephrotoxicity mechanisms, it is paramount to use new technologies, such as omic approaches (transcriptomics, proteomic, and metabolomics), bioinformatics platforms, CRISPR/Cas9 genome-editing and *in silico* modeling to advance this process. As shown in Figure. 3, these new technologies have been increasingly applied to study renal toxicity.

Omics technologies

Omics technologies include transcriptomics, proteomics, and metabolomics and are methodologies applied for biomarker identification useful in understanding the mechanisms of nephrotoxicity. Such technologies offer unbiased approaches to identify new biomarkers of AKI. Circulating or urinary microRNAs are being evaluated, and kidney-specific DNA methylation patterns are also being analyzed. Researchers have also been paying attention to urinary extra cell vesicles, which contain mRNA, microRNA, and proteins that mimic the cell physiology and pathophysiology along the nephron. The components of urinary extracellular vesicles are a source of information about disease pathogenesis, and they may serve as diagnostic and prognostic biomarkers, since they may enable target identification for drug discovery or provide evidence of biological activity in response to therapeutic agents. Among other objectives, omics studies seek to identify those that are mainly descriptive in nature, through the integration of experimental approaches and new computational modeling, such as in systems genetics, which will be

generated in order to help identify marker sets to guide diagnoses, monitor disease progression, and identify new therapeutic targets [8,20,106].

The European Consortium for High Throughput Research in Rare Kidney Diseases (EUREnOmics) has been working on a cohesive bioinformatics platform to study rare nephropathies. Additionally, a renal phenomenon database is being created, using the Human Phenotype Ontology website. The phenotype information will be linked to genomic, transcriptomic, proteomic, and metabolomics studies, omics datasets, and the public domain knowledge-base in a systems biology approach to identify molecular pathways associated with phenotypic features [107].

This situation thus becomes a great opportunity to advance *in vitro* studies for renal disease by analyzing gene expression. A bioinformatics platform will be an important tool to identify biomarkers that could be paramount in developing new medicines as well as in identifying nephrotoxicity and stopping the processes of AKI and CKD. Bioinformatics platforms are able to respond to such challenges as renal disease.

Transcriptomics: Studies on transcriptional profiling have revealed molecular markers and potential regulatory pathways of renal repair. A few key markers become active on the developmental pathways that have been reported during repair. By using high-resolution technologies, such as RNA sequencing and translational profiling specific to cell compartments within the kidney, further insights into injury and repair processes will be gained through the study of repairing transcriptome and cell-specific translome. An enhanced understanding holds promise for both the identification of modern therapeutic targets and biomarker-based evaluation of the damage-repair process [20].

Both proximal and distal tubular epithelial cells mount an acute transcriptional response to IRI. The earliest genes to be induced after *in vivo* injury (within 4 h after injury) include Fos, Jun, and Egr1 [108]. Fos is induced predominantly in the thick ascending limb (TAL) [109]. The latter observation suggests that the distal tubule, in addition to the proximal tubule, also senses the acute insertion. Subsequent microarray-based gene expression profiling studies encountered a similar immediate-early response, including Fos and Egr1, after IRI [110,111]. It is possible to identify toxicity on proximal and distal tubules, which will create an enormous advantage. Consequently, gene expression studies using tubular and renal cells may help to identify earlier nephrotoxicity [20].

Adler, *et al.* [112] showed that primary human proximal tubular epithelial cells (HPTEC) are a good *in vitro* model for screening for renal toxicity because they have differentiated epithelial cell characteristics. In addition, HO-1 expression proved to be a better option than currently used to predict toxicity in HPTEC, especially when combined with the total number of cells in the same assay. In addition, the approach used by the authors can be used to screen a large number of compounds besides can be used in combination with existing cytotoxicity assays.

The understanding of how the several cell types in the kidney communicate, in order to regulate the intrinsic repair mechanisms and their contribution to post injury fibrosis, remains a substantial challenge. Nonetheless, a continued focus is still placed on cell-type-specific signatures and a broadening of analysis beyond early injury, which in turn makes it feasible to provide important new insights. Clearly, the major clinical objective is to develop new analytical tools

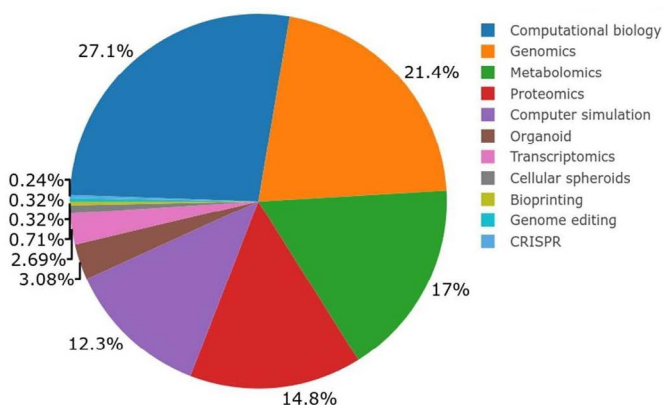


Figure 3. New technologies used to study renal toxicity

to diagnose both short- and long-term outcomes, as well as to develop new therapeutic strategies to improve existing repair processes and reduce the long-term risk of CKD after AKI [103]. In sum, all of these strategies could help to develop new agents to treat renal disease.

Proteomics: Proteomics is considered one of the most promising 'omics techniques for biomarker discovery in AKI, as it may identify specific molecular targets for early injury in different renal compartments. Almost all human diseases are characterized by a complex panorama at the molecular level, and it is imperative to acquire a global proteome picture to depict pathways and proteins with pivotal roles in pathogenesis. Proteomics studies have indicated that a single biomarker cannot fully account for the complexity of human diseases, and thus it is preferential to use biomarker panels.

Hence, it is thought that proteomics is a good tool for discovering and identifying new premature biomarkers. The main target of proteomics is to characterize the information flow within the cell and the organism, through protein pathways, interactions and networks, thereby allowing a hypothesis-free identification of disease biomarkers [113]. The search for biomarkers using urinary proteomics approaches has proven to be effective in CKD, diabetic nephropathy, AKI, ureteropelvic junction obstruction, vesicoureteral reflux, renal Fanconi syndrome, acute renal allograft rejection, cancer, DN, chronic rhinosinusitis, atherosclerosis and calcific aortic valve disease, and posterior urethral valves [106,114-119].

Metabolomics: Metabolomics is defined as the analysis of molecules smaller than 1,000 Da, which are transformed as a result of, and in support of, an organism's metabolism. The metabolome accounts for a complete set of metabolites that can be produced and consumed by organisms [117,120]. The results of metabolomics experiments constitute big data and require sophisticated data handling strategies and advanced statistical tools. In addition, before any statistical analysis, data preprocessing must take place to ensure the best possible results from the applied statistics [106]. Spectral alignment, normalization, transformation, and scaling are indispensable stages in metabolomic data preprocessing. Data normalization is required for accurate metabolite quantification. To normalize data, a common method is based on endogenous metabolites, such as urinary creatinine. However, creatinine itself may be somewhat variable, and subject to variation owing to diseases such as AKI [106,120,121].

The potential to identify drug targets from metabolomics will be amplified by the coordination of efforts to analyze samples of human kidney tissue and other biomaterial, such as urine and blood, by using state-of-the-art genomic, proteomic, and metabolomic approaches, together with detailed patient phenotyping and the use of existing biomarkers to discover and qualify new therapeutic purposes. Genetic data should be linked to existing phenotypic information or generated from customized human tissue models, by using induced pluripotent stem cells and targeted mutation, followed by differentiation to human kidney tissue. To give support to these efforts, enhanced disease prototypes are needed to reflect the complexity of CKD [117,122].

Equally challenging is the multifactorial origin of AKI and the fact that the changes of molecular expression induced by AKI are difficult to distinguish from those of the diseases associated with or causing AKI, such as shock or sepsis [106]. In the past, these experiments were focused on a single metabolite that was attributed to a specific disease or enzymatic reaction [123]. Nowadays, such technologies as GC/MS (gas chromatography-mass spectrometry) and computational tools allow for a more extensive and wide-ranging investigation of many

metabolites within a single measurement, providing a broader insight into mechanisms of diseases [106]. For this reason, it is very important to employ a panel of biomarkers.

New and cutting-edge techniques have been, and continue to be, developed; the challenge with the present status of metabolomics is its transition to clinical use. Integrating metabolomic data with other omics data for the purpose of drug discovery, and development is one direction in which this field may proceed [124,125].

The low sensitivity of current clinical markers (sCr and BUN) in early stages of the development of acute kidney injury limits their utility. Rapid LC/MS (liquid chromatography-mass spectrometry)-based metabolic profiling of serum was demonstrated in a key study showing that metabolomics could provide new indicators of AKI. Currently, omics technologies provide a great opportunity for research in rare renal diseases, as it is possible to probe the diseased organ directly [126]. Hence, we believe it is important to compare results obtained from patients and to evaluate them using cells, mainly because there are several different types of renal disease. This would thus be a safer strategy for using clinical results as a starting point.

CRISPR/Cas9 genome-editing

The fast advancement of genome-editing technologies has provided new pathways in the study of the genetic basis of human diseases and in the development of targeted therapeutic strategies that would not have been possible with traditional pharmacological drugs. Chemical toxicants act by many different mechanisms, however, the genes involved in adverse outcome pathways (AOPs) and AOP networks are not yet characterized. Functional genomic approaches can reveal both toxicity pathways and susceptibility genes, through knockdown or knockout of all non-essential genes in a cell of interest, and the identification of genes associated with a toxicity phenotype following toxicant exposure [127-129].

In this environment, new strategies could improve the engineering of genome-editing tools, and appropriate regulatory practices, together with genome editing, will accelerate discoveries in basic science and clinical translation, which will in turn aid in the process of identifying new biomarkers to become more robust. Hence, advances in genome editing with CRISPR/Cas9 will promote the use of kidney organoids to study inherited genetic kidney diseases [127].

The potential use of kidney organoids derived from hPSC combined with genome engineering technologies, particularly the CRISPR/Cas9, is a novel approach for the targeted modification of the renal epigenome to study renal organ dysfunction, such as CKD [130]. In addition, this technology could be further used to produce kidney phenotypes.

In silico

In silico, or computational, approaches to predicting and identifying compounds associated with kidney disease are based on the use of existing data and the ability to draw chemistry-based inferences from those data [131]. The concepts are relatively trivial in that, should data of acceptable quality be available for a compound, then there is no need for further testing. If no suitable data are available for a compound, then inferences can be made from it, eg, making an analogy with similar compounds or identifying chemical/molecular features that may be indicative of the ability to cause disease. All of these resources together form a battery of computational approaches, or even of a compound, that can shed light on the disease-causing

potential [132]. The general availability of data related to the ability of a compound to cause damage to the kidney has been reviewed by Pletz, *et al.* [133]. The resources can be classified according to the type of data they represent. The most relevant for human risk assessment are those for the effects of compounds in humans; these are available as part of clinical trial data or from reports of adverse drug reactions. Websites, such as www.clinicaltrials.gov, are useful to retrieve such data. While potentially relevant, they can only be used for the drugs to which they relate and lack consistency between clinical trials. To develop and evaluate *in silico* models, the results of animal tests done in the past are now useful, particularly low-dose chronic toxicity testing that show the superior predicting power of *in silico* models. Several resources have been compiled of repeated dose toxicity values and are freely available. These become useful in understanding and evaluating effects to the kidney when they record organ-level effects. The COSMOS database was developed specifically to achieve this and can be searched for organ-level toxicity, including effects to the kidney and bladder [134]. Other similar databases of repeated-dose toxicity data include REPDOSE [135] and HESS [136]. Other significant data resources of repeated-dose toxicity values that may require licensing and/or payment include Chemtunes (MN-AM, <https://www.mn-am.com/products/chemtunes>), Vitic Nexus (Lhasa Ltd, <https://www.lhasalimited.org/products/vitic-nexus.htm>), and LeadScope Toxicity Database (Leadscope Inc, http://www.leadscope.com/product_info.php?products_id=78). It is important to note that while these data resources may shed light on which compounds may promote disease, or damage, to the kidney, they were not developed specifically for this purpose.

There are fewer databases for *in vitro* data. However, one important resource that will aid in understanding, from the mechanistic point of view, is the compilation of the United States Tox21 initiative. Tox21 set about to screen large numbers of compounds (currently over 10,000) with high content assays, some of which are relevant to the effects and mechanisms of kidney disease and damage [137]. This immense and freely accessible resource is available through the US EPA's Chemistry Dashboard [138]. The Chemistry Dashboard is a cutting-edge and rapidly evolving resource for *in silico* assessment. It enables a user to input a chemical structure, retrieve existing (*in vivo* and Tox21) data, make predictions, and build knowledge about a substance. It is currently focused on a number of environmental and human health effects and could be applied to provide information on kidney disease if used appropriately, i.e., the data can assist in the identification and rationalization of chemicals associated with disease. Another important *in silico* tool that brings together data through the development of chemistry-based groups is the OECD (Organization for Economic Co-operation and Development) QSAR (Quantitative Structure-Activity Relationship) Toolbox (<https://www.qsartoolbox.org/>). Designed for grouping and read-across to allow for data gap filling, the Toolbox includes access to mechanistically based profilers to find similar molecules.

One of the most powerful methods to use existing data to make predictions for compounds with no data (the aforementioned process of data gap filling) is to form groups of similar compounds and "read-across" data from one molecule to another [139]. The use of the read-across paradigm has seen growth in recent years, most notably due to its application in the European Union's Registration, Evaluation, Authorization and restriction of Chemicals (REACH) legislation [140]. The identification of similar compounds is a broadly applied concept and requires a certain basis. One approach has been to use markers of organ-level toxicity to define similarity [141]. What seems particularly

relevant here are the chemistry-based structural alerts or biomarkers for kidney disease. Thus, should compounds be associated with the same chemistry alerts or biomarkers, then similarity in activity may be assumed, i.e., a prediction of harm can be made by analogy to the similar substance. While conceptually simple, there are numerous caveats to the approach and care must be taken [142]. One means of providing relevance from the *in silico* side is to ensure that the alerts are anchored on a mechanistic basis and specifically to relevant AOPs (see section below).

The logical identification of structure-activity relationships related to kidney disease, i.e., the sub-molecular features intimately associated with damage, has enabled the concepts of structural alerts to be defined. Such alerts are easily coded into software. Some structural alerts for kidney damage are available in the literature; for example, Myshkin, *et al.* [143] published alerts for a number of effects. Efforts are currently underway to rationalize the small number of publicly available structural alerts for harm caused to the kidney and to add new alerts where they are missing. This work is being supported by a compilation of over 200 compounds known to cause harm to the kidney [133], which have assisted in the evaluation of existing alerts and will enable efficient screening of new molecules for these effects.

Historically there have been many efforts to form quantitative associations between chemical structure and biological activity – the development of the so-called QSARs [131]. It is fundamental that QSARs require suitable data for development and, ideally, should be mechanistically based. As a result, particularly of the need for robust data sets, there have been relatively few QSAR models developed for nephrotoxicity; available models have been reviewed by Pletz, *et al.* [133]. For instance, Lei, *et al.* [144] used machine learning approaches to develop models for damage to the urinary tract. While some progress has been made, the development of QSARs for the full suite of harmful effects to the kidney is limited at the current time, but it can be expected to improve due to the better access to reliable data for modeling.

Using HPTEC and a set of 44 compounds, subdivided into HPTEC-toxic, non-HPTEC-toxic and non-nephrotoxic compounds, Su, *et al.* [145] described an approach based on high-throughput imaging, phenotypic profiling and machine learning methods to predict human nephrotoxicity. This approach does not require characterization of injury mechanisms and is capable of identifying structurally distinct xenobiotic-induced nephrotoxicity that lead similar phenotypic endpoints as from six sets of nuclear and actin cytoskeletal features with high efficiency and accuracy.

In addition to using information from chemistry for modeling, according to Soo, *et al.* [146], integration of information obtained from *in vitro* models into computational algorithms that incorporate patient-specific physiological parameters will not only reduce late drug loss from the development pipeline, but also facilitate the development of safer drugs and improve compound management of clinically important adverse nephrotoxic effects.

Stem cells

Stem cell therapy is an innovative approach to ameliorate IRI due to its antioxidative, immunomodulatory, and anti-apoptotic properties. However, the majority of the studies are confined to experimental animal models, although, several studies have investigated mesenchymal stem cells-based therapies for both acute and chronic kidney disease in clinical trials. Hence, it is paramount to understand the biological effects and mechanisms of action of stem cell therapy to improve its therapeutic benefits. In addition, more translational studies are needed to provide a more comprehensive understanding of

stem cell-based therapies and to ensure their safety for future clinical applications [147,148].

New models

New models such as organoids and 3D bioprinting also are important tools which may help us to understand nephrotoxicity mechanisms.

Organoids

Nephrons are complex architectural 3D structures. Hence, to recapitulate these structures *ex vivo*, the development of 3D culture systems is required. Organoids are 3D organ-like tissues that mimic *in vivo* organs structurally and functionally in culture plates. Kidney organoids derived from hPSCs are an attractive approach for studying mechanisms of human inherited kidney diseases, which might be applied to more common diseases as well as to the development of new drugs using human tissues, which may facilitate the translation of the findings obtained with this methodology to humans. However, to optimize this approach, differentiation protocols, genetic background, and epigenetic variation need to be considered in depth when disease phenotypes are analyzed in kidney organoids [99].

One goal of hPSC studies is to regenerate the kidney function. Kidneys form complex structures with blood filtration, and the urine re-absorption unit is paramount for their functioning and homeostasis. There are many challenges in the use of organoids to generate functional bioengineered kidney tissues. Vascularization of kidney organoids needs to be induced in an organized way to direct blood flow from arteries in order to drain into venous structures. This represents one of the great challenges related to vascularization [99].

3D bioprinting

In addition, the use of 3D bioprinting has proven to be an option to study nephrotoxicity, namely a functional bioengineered kidney with vasculature, multiple cell types, which include proximal tubular cells, and an extracellular matrix. A new system is being used to generate heterogeneous structures in 3D that enable the perfusion of liquid into lumens lined with vascular or tubular cells. It may be employed to mimic blood flow and intratubular flow in vascular and tubular channels [99,149,150].

Adverse outcome pathways

The AOP framework is a systematic, transparent approach used to organize existing toxicological knowledge and translate mechanistic information to adverse effects at higher levels of organization based on causal relationships between endpoints (Figure. 4). As such, the AOP framework has been proposed as a means to relate alternative types of data (*in silico*, *in vitro* and *in vivo*, biomarker-type data) to endpoints of concern to chemical risk assessors (eg, survival, growth, and reproduction). This application in the regulatory field is timely given

the current momentum to move toxicology from broadly empirical *in vivo* assessments of unique chemical effects to predictive approaches that make use of newer tools and alternative data [151-153].

With reference to AOPs for renal disease, a small number of studies have been performed using molecular initiating events (MIEs) to identify the following AO: adenomas and carcinomas, and renal failure and mortality, respectively. Most authors do not inform about the progress of the studies (<https://aopwiki.org/aops>).

Hence, it is paramount to use new tools in order to develop renal AOPs. Some studies have been performed utilizing biochemistry [154-156] and genomics [55-57]. However, it is important to make these data public through database, like the Comparative Toxicogenomics Database (CTD), in an attempt to achieve faster and more reliable results.

Thus, in terms of using the biomarker data, the AOP paradigm is a useful starting point. There are many advantages to the use of AOPs, including the rationalization of response and the ability to form the basis of quantitative AOPs, dissemination through the AOP-wiki, as well as the elucidation of the MIE [141]. In addition, the use of the MIE of an AOP is a robust means of developing structural alerts and has been applied for other organ-level toxicities [141,157]. At the time of writing this chapter, structural alerts for nephrotoxicity that allow for grouping and read-across are still under development [133].

It is clear that AOP will greatly help reduce animals in research involving nephrotoxicity, but it also uncovers the molecular mechanisms that may speed up the discovery of biomarkers for disease staging and therapy individualization, as well as the design and testing of novel therapeutic strategies that are safer to be developed, together with diagnoses and more accurate treatments.

Conclusion

The failure of existing animal models to accurately predict nephrotoxicity has been a major barrier to the development of safer drugs. The failure has mainly been due to the complexity of the physiology of the kidney, coupled with the variety and mechanisms of potential adverse effects, which animals are not able to identify. The application of advances in the understanding of molecular signaling and function would be a great benefit to find a means through which to prevent toxicity and/or treat kidney injury. Thus, new tools have been implemented, using alternative methods in order to elucidate mechanisms of nephrotoxicity. The use of *in vitro* models is at the forefront of the understanding of functional maturity and injury responses caused by drugs. However, it is essential to integrate existing *in vitro* methods with new technologies. For instance, organoids, 3D bioprinting, omics (transcriptomics, proteomics, and metabolomics), bioinformatics platforms, CRISPR/Cas9 genome-editing, and *in silico* modeling have been proposed as potential new tools to understand and identify nephrotoxicants. The main goal of the use of new and

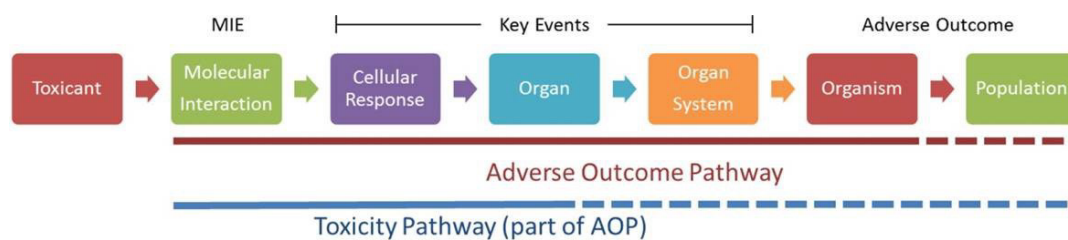


Figure 4. Adverse outcome pathways. Adapted from Browne P, Noyes PD, Casey WM, Dix DJ (2017) Application of Adverse Outcome Pathways to U.S. EPA's Endocrine Disruptor Screening Program. *Environ Health Perspect* 125: 096001.

existing technologies will be to generate an integrated model in which new biomarker(s) and/or pathway(s) involved in nephrotoxicity can be established as early as possible. The use of AOPs is helping to organize the information from the new technologies and the Molecular Initiating Event is expected to represent a robust means to detect toxicity. Therefore, combining a panel of biomarkers may improve the performance of the interpretation and predictability of nephrotoxicity. Gene expression integrated with computational analysis is also a key area that will facilitate the understanding of nephrotoxicity mechanisms and will consequently aid in the development of safer new drugs and be useful in the detection and monitoring of adverse effects from nephrotoxicants.

Acknowledgement

We thank the institutions that promote Brazilian research National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), Research Support Foundation of Minas Gerais State (Fapemig) and Pró-Reitoria de Pesquisa da UFMG (PRPq/UFMG).

Funding

Humane Society International

Disclosure

The author reports no conflicts of interest in this work.

References

- MacLeod A (2009) NCEPOD report on acute kidney injury—must do better. *Lancet* 374: 1405-1406.
- Kellum JA, Ronco C, Bellomo R (2016) Acute kidney disease and the community. *Lancet* 387: 1974-1976.
- Kramann R, Kusaba T, Humphreys BD (2015) Who regenerates the kidney tubule? *Nephrol Dial Transplant* 30: 903-910.
- Levin A, Tonelli M, Bonventre J, Coresh J, Donner JA, et al. (2017) Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy. *Lancet* 390: 1888-1917.
- KDIGO CKD Work Group (2013) KDIGO 2012 Clinical practice guideline for evaluation & management of CKD. *Kidney Int* 3: 5-14.
- Murugan R, Karajala-Subramanyam V, Lee M, Yende S, Kong L, et al. (2010) Acute kidney injury in non-severe pneumonia is associated with an increased immune response and lower survival. *Kidney Int* 77: 527-535.
- Mehta RL, Cerdá J, Burdmann EA, Tonelli M, García-García G, et al. (2015). International Society of Nephrology's Oby25 initiative for acute kidney injury (zero preventable deaths by 2025): a human rights case for nephrology. *Lancet* 385: 2616-2643.
- Ho J, Dart A, Rigatto C (2014) Proteomics in acute kidney injury - Current status and future promise. *Pediatr Nephrol* 29:163-171.
- Fortrie G, De Geus HRH, Betjes MGH (2019) The aftermath of acute kidney injury: A narrative review of long-term mortality and renal function. *Crit Care* 23: 1-11.
- Schmitt R, Coca S, Kanbay M, Tinetti ME, Cantley LG, et al. (2008) Recovery of kidney function after acute kidney injury in the elderly: A systematic review and meta-analysis. *Am J Kidney Dis* 52: 262-271.
- Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW (2005) Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. *J Am Soc Nephrol* 16: 3365-3370.
- Liyanaage T, Ninomiya T, Jha V, Neal B, Patrice HM, et al. (2015) Worldwide access to treatment for end-stage kidney disease: A systematic review. *Lancet* 385: 1975-1982.
- Bellomo R (2006) The epidemiology of acute renal failure: 1975 Versus 2005. *Curr Opin Crit Care* 12: 557-560.
- Kaufman J, Dhakal M, Patel B, Hamburger R (1991). Community-acquired acute renal failure. *Am J Kidney Dis* 17: 191-198.
- Nash K, Hafeez A, Hou S (2002) Hospital-acquired renal insufficiency. *Am J Kidney Dis* 39: 930-936.
- Wonnacott A, Meran S, Amphlett B, Talabani B, Phillips A (2014) Epidemiology and outcomes in community-acquired versus hospital-acquired aki. *Clin J Am Soc Nephrol* 9:1007-1014.
- Frost, Sullivan Rang HP (2007) Drug discovery and development. Churchill Livingstone: Elsevier.
- Devarajan P (2011) Biomarkers for the early detection of acute kidney injury. *Curr Opin Pediatr* 23: 194-200.
- Fuchs TC, Hewitt P (2011) Biomarkers for drug-induced renal damage and nephrotoxicity—An overview for applied toxicology. *AAPS J* 13: 615-631.
- Kumar S, Liu J, McMahon AP (2014) Defining the renal repair transcriptome after acute kidney injury. *Semin Nephrol* 34: 404-417.
- Lameire NH, Bagga A, Cruz D, De Maeseeneer J, Endre Z, et al. (2013) Acute kidney injury: an increasing global concern. *Lancet* 382: 170-179.
- Sales GTM, Foresto RD (2020) Drug-induced nephrotoxicity. *Rev Assoc Med Bras* 66: s82-s90.
- Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, et al. (2013) Chronic kidney disease: global dimension and perspectives. *Lancet* 382: 260-272.
- Horton R, Berman P (2015) Eliminating acute kidney injury by 2025: an achievable goal. *Lancet* 385: 2551-2552.
- Kashani K, Kellum JA (2015) Novel biomarkers indicating repair or progression after acute kidney injury. *Curr Opin Nephrol Hypertens* 24: 21-27.
- Zuk A, Bonventre JV (2016) Acute kidney injury. *Annu Rev Med* 67: 293-307.
- Waheed S, Matsushita K, Astor BC, Hoogeveen RC, Ballantyne C, et al. (2013) Combined association of creatinine, albuminuria, and cystatin C with all-cause mortality and cardiovascular and kidney outcomes. *Clin J Am Soc Nephrol* 8: 434-442.
- Stephen C, Noble J, Simpson K, Allison MEM (2006) Long-term quality of life and hospital mortality in patients treated with intermittent or continuous hemodialysis for acute renal and respiratory failure. *Ren Fail* 28: 323-330.
- Balogun SA, Balogun R, Philbrick J, Abdel-Rahman E (2017) Quality of life, perceptions, and health satisfaction of older adults with end-stage renal disease: A systematic review. *J Am Geriatr Soc* 65: 777-785.
- Eckardt K-U, Coresh J, Devuyst O, Johnson RJ, Köttgen A, et al. (2013) Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet* 382: 158-169.
- Cazorla-Vázquez S, Engel FB (2018) Adhesion GPCRs in kidney development and disease. *Front Cell Dev Biol* 6: 9.
- Pickkers P, Ostermann M, Joannidis M, Zarbock A, Hoste E, et al. (2017) The intensive care medicine agenda on acute kidney injury. *Intensive Care Med* 43: 1198-1209.
- Muralidharan A, White S (2015) The need for kidney transplantation in low- and middle-income countries in 2012. *Transplantation* 99: 476-481.
- Luyckx VA, Tonelli M, Stanifer JW (2018) The global burden of kidney disease and the sustainable development goals. *Bull World Health Organ* 96: 414-422.
- Norton JM, Moxey-Mims MM, Eggers PW, Narva AS, Star RA, et al. (2016) Social determinants of racial disparities in CKD. *J Am Soc Nephrol* 27: 2576-2595.
- Ojo A (2014) Addressing the global burden of chronic kidney disease through clinical and translational research. *Trans Am Clin Climatol Assoc* 125: 229-246.
- Sanchez-Niño MD, Sanz AB, Ramos AM, Ruiz-Ortega M, Ortiz A (2017) Translational science in chronic kidney disease. *Clin Sci* 131: 1617-1629.
- Kam-Tao Li P, Burdmann EA, Mehta RL (2013) Acute kidney injury: global health alert. *J Nephropathol* 2: 90-97.
- Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, et al. (2012) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the global burden of disease study 2010. *Lancet* 380: 2095-2128.
- Humphreys BD (2018) Mechanisms of renal fibrosis. *Annu Rev Physiol* 80: 309-326.
- Liu Y (2011) Cellular and molecular mechanisms of renal fibrosis. *Nat Rev Nephrol* 7: 684-696.
- Rao C, Adair T, Bain C, Doi SAR (2012) Mortality from diabetic renal disease: A hidden epidemic. *Eur J Public Health* 22: 280-284.

43. Pickering JW, Endre ZH (2016) Bench to bedside: The next steps for biomarkers in acute kidney injury. *Am J Physiol - Ren Physiol* 311: F717-F721.
44. Alge JL, Arthur JM (2015) Biomarkers of AKI: A review of mechanistic relevance and potential therapeutic implications. *Clin J Am Soc Nephrol* 10: 147-155.
45. Koyner JL, Garg AX, Coca SG, Sint K, Thiessen-Philbrook H, et al. (2012) Biomarkers predict progression of acute kidney injury after cardiac surgery. *J Am Soc Nephrol* 23: 905-914.
46. Meersch M, Volmering S, Zarbock A (2017) Prevention of acute kidney injury. *Best Pract Res Clin Anaesthesiol* 31: 361-370.
47. Schrezenmeier EV, Barasch J, Budde K, Westhoff T, Schmidt-Ott KM (2017) Biomarkers in acute kidney injury - pathophysiological basis and clinical performance. *Acta Physiol* 219: 556-574.
48. Block CA, Schoolwerth AC (2006) Critical care issues for the nephrologist: The epidemiology and outcome of acute renal failure and the impact on chronic kidney disease. *Semin Dial* 19: 450-454.
49. Devarajan P (2008) The future of pediatric acute kidney injury management-biomarkers. *Semin Nephrol* 28: 493-498.
50. Hillege HL, Nitsch D, Pfeffer MA, Swedberg K, McMurray JJV, et al. (2006) Renal function as a predictor of outcome in a broad spectrum of patients with heart failure. *Circulation* 113: 671-678.
51. Ichimura T, Hung CC, Yang SA, Stevens JL, Bonventre JV (2004) Kidney injury molecule-1: a tissue and urinary biomarker for nephrotoxicant-induced renal injury. *Am J Physiol Physiol* 286: F552-F563.
52. Medić B, Rovčanin B, Basta Jovanović G, Radojević-Škodrić S, Prostran M (2015) Kidney injury molecule-1 and cardiovascular diseases: From basic science to clinical practice. *Biomed Res Int* 2015: 1-10.
53. Miao J, Friedman E, Wu AHB, Todd JA, Estis J, et al. (2017) Clinical utility of single molecule counting technology for quantification of KIM-1 in patients with heart failure and chronic kidney disease. *Clin Biochem* 50: 889-895.
54. Ichimura T, Bonventre JV, Bailly V, Wei H, Hession CA, et al. (1998) Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion molecule containing a novel immunoglobulin domain, is up-regulated in renal cells after injury. *J Biol Chem* 273: 4135-4142.
55. Campos MAA, Almeida LA de, Grossi MF, Tagliati CA (2018) *In vitro* evaluation of biomarkers of nephrotoxicity through gene expression using gentamicin. *J Biochem Mol Toxicol* 32: e22189.
56. Silva SCT, de Almeida LA, Soares S, Grossi MF, Valente AMS, et al. (2017) *In vitro* study of putative genomic biomarkers of nephrotoxicity through differential gene expression using gentamicin. *Toxicol Mech Methods* 27: 435-441.
57. Grossi MF, Campos MAA, Soares S, Silva SCT, Nunes MS, et al. (2017) *In vitro* study of potential nephrotoxicity biomarkers through gene expression using Amphotericin B. *J Toxicol Pharmacol* 1: 1008.
58. Camacho P, Fan H, Liu Z, He JQ (2016) Small mammalian animal models of heart disease. *Am J Cardiovasc Dis* 3: 70-80.
59. Nath KA. 2015. Models of human AKI: Resemblance, reproducibility, and return on investment. *J Am Soc Nephrol* 26: 2891-2893.
60. Martin-Sanchez D, Ruiz-Andres O, Poveda J, Carrasco S, Cannata-Ortiz P, et al. (2017) Ferroptosis, but not necroptosis, is important in nephrotoxic folic acid-induced AKI. *J Am Soc Nephrol* 28: 218-229.
61. Ortiz A, Sanchez-Niño MD, Izquierdo MC, Martin-Cleary C, Garcia-Bermejo L, et al. (2015) Translational value of animal models of kidney failure. *Eur J Pharmacol* 759: 205-220.
62. Bao Y-W, Yuan Y, Chen J-H, Lin W-Q (2018) Kidney disease models: tools to identify mechanisms and potential therapeutic targets. *Zool Res* 39: 72-86.
63. Hayward RS, Harding J, Molloy R, Land L, Longcroft-Neal K, et al. (2018) Adverse effects of a single dose of gentamicin in adults: a systematic review. *Bri J Clin Pharmacol* 84: 223-238.
64. Lavergne A, Vigneau C, Polard E, Triquet L, Rioux-Leclercq N, et al. (2018) Acute kidney injury during treatment with high-dose cloxacillin: a report of 23 cases and literature review. *Int J Antimicrob Agents* 52: 344-349.
65. Quintanilha JCF, Saavedra KF, Visacri MB, Moriel P, Salazar LA (2019) Role of epigenetic mechanisms in cisplatin-induced toxicity. *Crit Rev Oncol Hematol* 137: 131-142.
66. Uccelli A, Moretta L, Pistoia V (2008) Mesenchymal stem cells in health and disease. *Nat Rev Immunol* 8: 726-736.
67. Vanherweghem J-L, Tielemans C, Abramowicz D, Depierreux M, Vanhaelen-Fastre R, et al. (1993) Rapidly progressive interstitial renal fibrosis in young women: association with slimming regimen including Chinese herbs. *Lancet* 341: 387-391.
68. <http://wayback.archiveit.org/7993/20171115154508/https://www.fda.gov/Food/RecallsOutbreaksEmergencies/SafetyAlertsAdvisories/ucm096388.htm>
69. Fink M, Henry M, Tange JD (1987) Experimental folic acid nephropathy. *Pathology* 19: 143-149.
70. Linkermann A, Skouta R, Himmerkus N, Mulay SR, Dewitz C, et al. (2014) Synchronized renal tubular cell death involves ferroptosis. *Proc Natl Acad Sci* 111: 16836-16841.
71. Linkermann A, Bräsen JH, Himmerkus N, Liu S, Huber TB, et al. (2012) Rip1 (Receptor-interacting protein kinase 1) mediates necroptosis and contributes to renal ischemia/reperfusion injury. *Kidney Int* 81: 751-761.
72. Khan KM, Jialal I (2018) Folic Acid (Folate) Deficiency. *StatPearls* 2018: 1.
73. Brodsky SV, Satoskar A, Chen J, Nadasdy G, Eagen JW, et al. (2009) Acute kidney injury during warfarin therapy associated with obstructive tubular red blood cell casts: A report of 9 Cases. *Am J Kidney Dis* 54: 1121-1126.
74. Brodsky SV, Nadasdy T, Rovin BH, Satoskar AA, Nadasdy GM, et al. (2011) Warfarin-related nephropathy occurs in patients with and without chronic kidney disease and is associated with an increased mortality rate. *Kidney Int* 80: 181-189.
75. Ware K, Brodsky P, Satoskar AA, Nadasdy T, Nadasdy G, et al. (2011) Warfarin-related nephropathy modeled by nephron reduction and excessive anticoagulation. *J Am Soc Nephrol* 22: 1856-1862.
76. Elterman J, Zonies D, Stewart I, Fang R, Schreiber M (2015) Rhabdomyolysis and acute kidney injury in the injured war fighter. *J Trauma Acute Care Surg* 79: S171-S174.
77. Zhang L, Kang Y, Fu P, Cao Y, Shi Y, et al. (2012) Myoglobin clearance by continuous venous-venous haemofiltration in rhabdomyolysis with acute kidney injury: A case series. *Injury* 43: 619-623.
78. Geng X, Wang Y, Hong Q, Yang J, Zheng W, et al. (2015) Differences in gene expression profiles and signaling pathways in rhabdomyolysis-induced acute kidney injury. *Int J Clin Exp Pathol* 8: 14087-14098.
79. Kim JH, Lee D-W, Jung MH, Cho H-S, Jeon D-H, et al. (2014) Macrophage depletion ameliorates glycerol-induced acute kidney injury in mice. *Nephron Exp Nephrol* 128: 21-29.
80. Singh P, Ricksten SE, Bragadottir G, Redfors B, Nordquist L (2013) Renal oxygenation and haemodynamics in acute kidney injury and chronic kidney disease. *Clin Exp Pharmacol Physiol* 40: 138-147.
81. Tampe B, Steinle U, Tampe D, Carstens JL, Korsten P, et al. (2017) Low-dose hydralazine prevents fibrosis in a murine model of acute kidney injury-to-chronic kidney disease progression. *Kidney Int* 91: 157-176.
82. Endre ZH, Pickering JW (2013) Biomarkers and creatinine in AKI: The trough of disillusionment or the slope of enlightenment? *Kidney Int* 84: 644-647.
83. Ergür BU, Mıcılı SÇ, Yılmaz O, Akokay P (2015) The effects of α -lipoic acid on aortic injury and hypertension in the rat remnant kidney (5/6 nephrectomy) model. *Anatol J Cardiol* 15: 443-449.
84. Tapia E, Soto V, Ortiz-Vega KM, Zarco-Márquez G, Molina-Jijón E, et al. (2012) Curcumin induces nrf2 nuclear translocation and prevents glomerular hypertension, hyperfiltration, oxidant stress, and the decrease in antioxidant enzymes in 5/6 nephrectomized rats. *Oxid Med Cell Longev* 2012: 1-14.
85. Deb DK, Sun T, Wong KE, Zhang Z, Ning G, et al. (2010) Combined vitamin D analog and AT1 receptor antagonist synergistically block the development of kidney disease in a model of type 2 diabetes. *Kidney Int* 77: 1000-1009.
86. Kitada M, Ogura Y, Koya D (2016) Rodent models of diabetic nephropathy: Their utility and limitations. *Int J Nephrol Renovasc Dis* 9: 279-290.
87. Kong L, Wu H, Cui W, Zhou W, Luo P, et al. (2013) Advances in murine models of diabetic nephropathy. *J Diabetes Res* 2013: 1-10.
88. Nishimura A, Kasai T, Kikuno S, Nagasawa K, Okubo M, et al. (2018) Effect of sleep-disordered breathing on albuminuria in 273 patients with Type 2 Diabetes. *J Clin Sleep Med* 14: 401-407.

89. Ruiz-Ortega M, Esteban V, Rupérez M, Sánchez-López E, Rodríguez-Vita J, et al. (2006) Renal and vascular hypertension-induced inflammation: role of angiotensin II. *Curr Opin Nephrol Hypertens* 15: 159-166.
90. Fogo AB (2015) Causes and pathogenesis of focal segmental glomerulosclerosis. *Nat Rev Nephrol* 11: 76-87.
91. Daehn I, Casalena G, Zhang T, Shi S, Fenninger F, et al. (2014) Endothelial mitochondrial oxidative stress determines podocyte depletion in segmental glomerulosclerosis. *J Clin Invest* 124: 1608-1621.
92. Joyce E, Glasner P, Ranganathan S, Swiatecka-Urban A (2017) Tubulointerstitial nephritis: diagnosis, treatment, and monitoring. *Pediatr Nephrol* 32: 577-587.
93. O’Kell AL, Grant DC, Khan SR (2017) Pathogenesis of calcium oxalate urinary stone disease: species comparison of humans, dogs, and cats. *Urolithiasis* 45: 329-336.
94. Fonseca KL, Rodrigues PNS, Olsson IAS, Saraiva M (2017) Experimental study of tuberculosis: From animal models to complex cell systems and organoids. *PLoS Pathogens* 40: e1006421.
95. Eisenbrand G, Pool-Zobel B, Baker V, Balls M, Blaauboer B, et al. (2002) Methods of *in vitro* toxicology. *Food Chem Toxicol* 40: 19-236.
96. Descotes J (2003) From clinical to human toxicology: linking animal research and risk assessment in man. *Toxicol Lett* 140-141: 3-10.
97. Bernauer U, Oberemm A, Madle S, Gundert-Remy U (2005) The Use of *in vitro* data in risk assessment. *Basic Clin Pharmacol & Toxicol* 96: 176-181.
98. Innovate UK (2015) A non-animal technologies roadmap for the UK: Advancing predictive biology.
99. Morizane R, Bonventre JV (2017) Kidney organoids: A translational journey. *Trends Mol Med* 23: 246-263.
100. Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, et al. (2014) A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515: 355-364.
101. Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, et al. (2014) Principles of regulatory information conservation between mouse and human. *Nature* 515: 371-375.
102. Dakhore S, Nayer B, Hasegawa K (2018) Human pluripotent stem cell culture: Current status, challenges, and advancement. *Stem Cells Int* 2018: 1-17.
103. Devuyst O, Knoers NVAM, Remuzzi G, Schaefer F (2014) Rare inherited kidney diseases: challenges, opportunities, and perspectives. *Lancet* 383: 1844-1859.
104. Stergachis AB, Neph S, Sandstrom R, Haugen E, Reynolds AP, et al. (2014) Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature* 515: 365-370.
105. PhRMA (2013) 2013 Biopharmaceutical research industry profile. *Biopharm Res Ind Phrma* 1: 78.
106. Marx D, Metzger J, Pejchinovski M, Gil RB, Frantzi M, et al. (2018) Proteomics and metabolomics for AKI diagnosis. *Semin Nephrol* 38: 63-87.
107. Levy SE, Myers RM (2016) Advancements in next-generation sequencing. *Annu Rev Genomics Hum Genet* 17: 95-115.
108. Ouellette AJ, Malt RA, Sukhatme VP, Bonventre JV (1990) Expression of two “immediate early” genes, Egr-1 and c-fos, in response to renal ischemia and during compensatory renal hypertrophy in mice. *J Clin Invest* 85: 766-771.
109. Witzgall R, Brown D, Schwarz C, Bonventre JV (1994) Localization of proliferating cell nuclear antigen, vimentin, c-Fos, and clusterin in the posts ischemic kidney. Evidence for a heterogenous genetic response among nephron segments, and a large pool of mitotically active and dedifferentiated cells. *J Clin Invest* 93: 2175-2188.
110. Supavekin S, Zhang W, Kucherlapati R, Kaskel FJ, Moore LC, et al. (2003) Differential gene expression following early renal ischemia/reperfusion. *Kidney Int* 63: 1714-1724.
111. Yuen PST, Jo SK, Holly MK, Hu X, Star RA (2006) Ischemic and nephrotoxic acute renal failure are distinguished by their broad transcriptomic responses. *Physiol Genomics* 25: 375-386.
112. Adler M, Ramm S, Hafner M, Muhlich JL, Gottwald EM, et al. (2016) A quantitative approach to screen for nephrotoxic compounds *in vitro*. *J Am Soc Nephrol* 27: 1015-1028.
113. Rosner MH, Okusa MD (2006) Acute kidney injury associated with cardiac surgery. *Clin J Am Soc Nephrol* 1: 19-32.
114. Di Carlo C, Brandi J, Cecconi D (2018) Pancreatic cancer stem cells: Perspectives on potential therapeutic approaches of pancreatic ductal adenocarcinoma. *World J Stem Cells* 10: 172-182.
115. Mulligan JK, Patel K, Williamson T, Reaves N, Carroll W, et al. (2018) C3a receptor antagonism as a novel therapeutic target for chronic rhinosinusitis. *Mucosal Immunol* 11: 1375-1385.
116. Neufeld EB, Zadrozny LM, Phillips D, Aponte A, Yu ZX, et al. (2014) Decorin and biglycan retain LDL in disease-prone valvular and aortic subendothelial intimal matrix. *Atherosclerosis* 233: 113-121.
117. Patti GJ, Yanes O, Siuzdak G (2012) Metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol* 13: 263-269.
118. Salvadori M, Tsalouchos A (2017) Biomarkers in renal transplantation: An updated review. *World J Transplant* 7: 161.
119. Van JAD, Scholey JW, Konvalinka A (2017) Insights into diabetic kidney disease using urinary proteomics and bioinformatics. *J Am Soc Nephrol* 28: 1050-1061.
120. Alonso A, Marsal S, Julià A (2015) Analytical methods in untargeted metabolomics: State of the art in 2015. *Front Bioeng Biotech* 3: 23.
121. Dieterle F, Ross A, Schlotterbeck G, Senn H (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabolomics. *Anal Chem* 78: 4281-4290.
122. Nicholson JK, Lindon JC (2008) Metabonomics. *Nature* 455: 1054-1056.
123. German JB, Hammock BD, Watkins SM (2005) Metabolomics: building on a century of biochemistry to guide human health. *Metabolomics* 1: 3-9.
124. Kell DB, Goodacre R (2014) Metabolomics and systems pharmacology: why and how to model the human metabolic network for drug discovery. *Drug Discov Today* 19: 171-182.
125. Robertson DG, Frevert U (2013) Metabolomics in drug discovery and development. *Clin Pharmacol Ther* 94: 559-561.
126. Sun J, Shannon M, Ando Y, Schnackenberg LK, Khan NA, et al. (2012) Serum metabolomic profiles from patients with acute kidney injury: A pilot study. *J Chromatogr B* 893-894: 107-113.
127. Lin J, Musunuru K (2016) Genome engineering tools for building cellular models of disease. *FEBS* 283: 3222-3231.
128. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, et al. (2014) Genome-Scale CRISPR-Cas9 knockout screening in human cells. *Science* 343: 84-87.
129. Shen H, McHale CM, Smith MT, Zhang L (2015) Functional genomic screening approaches in mechanistic toxicology and potential future applications of CRISPR-Cas9. *Mutat Res Mutat Res* 764: 31-42.
130. Hurtado Del Pozo C, Garreta E, Izpisua Belmonte JC, Montserrat N (2018) Modeling epigenetic modifications in renal development and disease with organoids and genome editing. *Dis Model Mech* 11: dmm035048.
131. Holtzman S (2006) In silico toxicology. *Ann N Y Acad Sci* 919: 68-74.
132. Cronin MTD, Yoon M (2018) Historical context of computational methods to predict toxicity including (Q)SAR, chemoinformatics, grouping, read-across and PBPK modelling. In: Balls M, Combes R, Worth AP, eds. *The History of Alternative Test Methods in Toxicology*. In press, Academic Press.
133. Pletz J, Enoch SJ, Jais DM, Mellor CL, Pawar G, et al. (2018) A critical review of adverse effects to the kidney: mechanisms, data sources, and in silico tools to assist prediction. *Expert Opin Drug Metab Toxicol* 14: 1225-1253.
134. Mostrag A, Yang C, Cronin M, Madden J, Schwab C, et al. (2018) Supporting data location, read-across, and toxicity data gap filling using the COSMOS database. *Toxicol* 162: 374-375.
135. Bitsch A, Jacobi S, Melber C, Wahnschaffe U, Simetska N, et al. (2006) REPOSE: A database on repeated dose toxicity studies of commercial chemicals—A multifunctional tool. *Regul Toxicol Pharmacol* 46: 202-210.
136. Sakuratani Y, Zhang HQ, Nishikawa S, Yamazaki K, Yamada T, et al. (2013) Hazard Evaluation Support System (HESS) for predicting repeated dose toxicity using toxicological categories. *SAR QSAR Environ Res* 24: 351-363.
137. Richard AM, Judson RS, Houck KA, Grulke CM, Volarath P, et al. (2016) ToxCast Chemical Landscape: Paving the road to 21st century toxicology. *Chem Res Toxicol* 29: 1225-1251.
138. Williams AJ, Grulke CM, Edwards J, McEachran AD, Mansouri K, et al. (2017) The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J Cheminform* 9: 61.

139. Patlewicz G, Cronin MTD, Helman G, Lambert JC, Lizarraga LE, et al. (2018) Navigating through the minefield of read-across frameworks: A commentary perspective. *Comput Toxicol* 6: 39-54.
140. Spielmann H, Sauer UG, Mekenyan O (2011) A critical evaluation of the 2011 ECHA reports on compliance with the REACH and CLP regulations and on the use of alternatives to testing on animals for compliance with the REACH regulation. *ATLA* 39: 481-493.
141. Cronin MTD, Richarz A-N (2017) Relationship between adverse outcome pathways and chemistry-based *in silico* models to predict toxicity. *Appl Vitro Toxicol* 3: 286-297.
142. Schultz TW, Cronin MTD (2017) Lessons learned from read-across case studies for repeated-dose toxicity. *Regul Toxicol Pharmacol* 88: 185-191.
143. Myshkin E, Brennan R, Khasanova T, Sitnik T, Serebriyskaya T, et al. (2012) Prediction of organ toxicity endpoints by qsar modeling based on precise chemical-histopathology annotations. *Chem Biol Drug Des* 80: 406-416.
144. Lei T, Sun H, Kang Y, Zhu F, Liu H, et al. (2017) ADMET evaluation in drug discovery. 18. Reliable prediction of chemical-induced urinary tract toxicity by boosting machine learning approaches. *Mol Pharm* 14: 3935-3953.
145. Su R, Xiong S, Zink D, Loo LH (2016) High-throughput imaging-based nephrotoxicity prediction for xenobiotics with diverse chemical structures. *Arch Toxicol* 90: 2793-2808.
146. Soo JY-C, Jansen J, Masereeuw R, Little MH (2018) Advances in predictive *in vitro* models of drug-induced nephrotoxicity. *Nat Rev Nephrol* 14: 378-393.
147. Lee K-H, Tseng W-C, Yang C-Y, Tarng D-C (2019) The Anti-inflammatory, anti-oxidative, and anti-apoptotic benefits of stem cells in acute ischemic kidney injury. *Int J Mol Sci* 20: 3529.
148. Li J-S, Li B (2019) Renal injury repair: How about the role of stem cells. In: Liu B, Lan H, Lv L, eds. Renal fibrosis: Mechanisms and therapies. Singapore: Springer. p. 661-670.
149. Kolesky DB, Homan KA, Skylar-Scott MA, Lewis JA (2016) Three-dimensional bioprinting of thick vascularized tissues. *Proc Natl Acad Sci* 113: 3179-3184.
150. Kolesky DB, Truby RL, Gladman AS, Busbee TA, Homan KA, et al. (2014) 3D bioprinting of vascularized, heterogeneous cell-laden tissue constructs. *Adv Mater* 26: 3124-3130.
151. Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, et al. (2010) Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem* 29: 730-741.
152. Fay KA, Villeneuve DL, LaLone CA, Song Y, Tollefsen KE, et al. (2017) Practical approaches to adverse outcome pathway development and weight-of-evidence evaluation as illustrated by ecotoxicological case studies. *Environ Toxicol Chem* 36: 1429-1449.
153. National Research Council (2007) Toxicity testing in the 21st century. Toxicity testing in the 21st century: A vision and a strategy. Washington DC: National Academies Press pp: 1-196.
154. Ferreira AF, França FD, Rossoni JV, Viana PHL, Moraes KCM, et al. (2016) Cyclic adenosine monophosphate protects renal cell lines against amphotericin B toxicity in a PKA-independent manner. *Drug Chem Toxicol* 39: 28-34.
155. França FD, Ferreira AF, Lara RC, Rossoni JV, Costa DC, et al. (2014) Role of protein kinase A signaling pathway in cyclosporine nephrotoxicity. *Toxicol Mech Methods* 24: 369-376.
156. França FD, Ferreira AF, Lara RC, Rossoni JV, Costa DC, et al. (2014) Alteration in cellular viability, pro-inflammatory cytokines and nitric oxide production in nephrotoxicity generation by Amphotericin B: involvement of PKA pathway signaling. *J Appl Toxicol* 34: 1285-1292.
157. Nelms MD, Mellor CL, Cronin MTD, Madden JC, Enoch SJ (2015) Development of an *in silico* profiler for mitochondrial toxicity. *Chem Res Toxicol* 28: 1891-1902.