

**Faculté de pharmacie
et des sciences biomédicales**

Caractérisation de longs ARN non-codants liés à l'épuisement des lymphocytes T CD8⁺ intra-tumoraux

Auteur : Florian Van Horenbeke

Promoteur : Pierre van der Bruggen

Lecteurs : Thomas Michiels & Benoît Van den Eynde

Année académique 2022-2023

Master en sciences biomédicales à finalité approfondie, option oncologie

Remerciements

Ce stage de mémoire était un voyage parsemé d'obstacles à surmonter, mais avec l'aide et le soutien des personnes qui m'entouraient, j'ai finalement atteint ma destination, la remise de ce mémoire.

Je tiens tout d'abord à remercier le professeur Pierre van der Bruggen pour m'avoir accueilli dans son laboratoire, pour sa disponibilité à toute épreuve, ses nombreuses relectures et son implication dans ma formation.

Un grand merci à Svetlana, mon encadrante, pour son accueil et son aide précieuse dans les premiers mois de ce stage, qui sont les plus importants. Merci pour ta bienveillance et ton enthousiasme communicatif pour les longs ARN non-codants.

Comment écrire des remerciements et ne pas remercier l'équipe TILS ? Pour moi, j'ai commencé ce stage avec une encadrante et je l'ai fini avec pas moins de huit ! Merci à vous pour m'avoir accueilli à bras ouverts et pour avoir pris le temps de m'encadrer au fil des expériences et des pauses café. Merci à Alexandre, Annika, Céline, Christophe, Claude, Nicolas, Shiqiang et Thibault.

Je souhaiterais aussi remercier les âmes généreuses qui ont donné de leur temps pour relire ce mémoire. Merci à Camille, Céline, Claude et Constance.

Je souhaiterais également remercier individuellement Axelle, sans qui cette année n'aurait pas eu la même saveur. Que ce soit nos rendez-vous quotidiens autour du café noisette, nos discussions intenses sur nos sujets de mémoire ou nos précieuses corrections mutuelles, cette fabuleuse amitié fait partie intégrante de cette aventure de vie qu'à été mon mémoire.

Je remercie aussi tous les hommes et les femmes de l'ombre sans qui rien de ce mémoire n'aurait-été possible. Merci à Nicolas Dauguet, Nora, Samira et Olivier.

Je remercie chaleureusement ma famille et mes amis qui m'ont soutenu tout au long de cette aventure.

Merci à toutes les personnes trop nombreuses que j'oublie de citer et qui m'ont aidé tout au long de ce stage.

Enfin, merci à vous qui lisez mon travail.

Résumé

Le système immunitaire est capable de reconnaître et d'éliminer les cellules cancéreuses. Néanmoins, une perte graduelle des fonctions effectrices des lymphocytes T est observée lorsqu'ils sont chroniquement activés par les cellules cancéreuses. Ce processus est appelé "l'épuisement" et pourrait favoriser le développement de nombreux cancers. Mieux comprendre ce phénomène apparaît comme une approche pertinente pour améliorer les traitements d'immunothérapie, tels que les thérapies adoptives de lymphocytes T génétiquement modifiés.

Le rôle des longs ARN non-codants (lncRNAs) n'a été que très peu étudié dans l'épuisement, en particulier chez l'Homme. Longtemps considérés comme un bruit de fond génomique sans fonction, on sait aujourd'hui que de nombreux lncRNAs jouent des rôles essentiels, notamment dans les fonctions effectrices des lymphocytes T CD8⁺.

Le premier objectif de ce mémoire était de mettre au point une approche bio-informatique afin d'identifier des lncRNAs potentiellement impliqués dans l'épuisement des lymphocytes T CD8⁺ intra-tumoraux. Nous avons mis au point une approche combinant l'analyse des lncRNAs différentiellement exprimés et l'analyse des réseaux de gènes co-exprimés. Cette analyse permet d'identifier des lncRNAs liés à un phénomène d'intérêt, tout en caractérisant leurs fonctions potentielles par l'analyse des gènes codants qui y sont co-exprimés. Au terme de cette analyse, une dizaine de lncRNAs ont été identifiés. Parmi eux, LINC00402 a été choisi pour être étudié plus en profondeur.

L'objectif secondaire de mon mémoire était de mettre au point une méthode de caractérisation fonctionnelle de ces lncRNAs. L'électroporation de petits ARN interférents a été testée dans des lymphocytes T CD8⁺ et s'est révélée décevante. Nous nous sommes alors tournés vers l'utilisation de GapmeR. Des expériences préliminaires ont montré que l'utilisation d'un GapmeR ciblant LINC00402 a un effet sur la prolifération et la production d'IFN γ et de TNF α des lymphocytes T CD8⁺.

Abstract

The immune system is able to recognize and eliminate cancer cells. Nonetheless, when T lymphocytes are chronically activated by cancer cells, they gradually lose their effector functions. This phenomenon is usually referred to as "exhaustion" and could explain the proliferation of the cancer cells that escape immune control. Understanding this phenomenon seems to be a relevant approach to improve immunotherapy treatments, such as T-cell transfer therapies.

The role of long non-coding RNAs (lncRNAs) in exhausted T cells has received very little attention, particularly in humans. Previously considered as useless genomic noise, many lncRNAs are now known to play essential roles, particularly in CD8⁺ T cell effector functions.

The first objective of my thesis was to develop a bioinformatic approach to identify lncRNAs potentially involved in the exhaustion of intratumoral CD8⁺ T lymphocytes. We developed an approach combining the analysis of the differentially expressed lncRNAs between two conditions with the analysis of co-expressed gene networks. This analysis allows the identification of lncRNAs related to a phenomenon of interest while characterizing their potential functions by analyzing their known co-expressed coding genes. At the end of this analysis, a dozen lncRNAs potentially involved in exhaustion were identified. Among them, LINC00402 was chosen for in-depth characterization.

The second objective of my thesis was to develop a functional characterization method for these lncRNAs. The electroporation of small interfering RNAs was tested in CD8⁺ T lymphocytes and proved disappointing. We used GapmeRs. Preliminary experiments showed that the use of a GapmeR targeting LINC00402 has an effect on proliferation and the production of IFN γ and TNF α of CD8⁺ T lymphocytes.

Table des matières

Introduction	5
1. Une brève histoire de l'épuisement des lymphocytes intra-tumoraux	5
1.1. Découverte de l'épuisement : Le modèle LCMV	6
1.2. Existence de l'épuisement dans les tumeurs humaines et murines.	7
1.3. Identifier les lymphocytes T épuisés grâce au facteur de transcription TOX	8
2. Les longs ARN non-codants	10
3. Challenges et objectifs	13
4. Stratégie mise en place pour accomplir nos objectifs	15
Résultats	16
1. Aperçu général de mes données	16
2. Définition des lncRNAs différentiellement exprimés	17
3. Analyse des réseaux de co-expression génique	20
3.1. Identification des modules de co-expressions	20
3.2. Prédictions fonctionnelles de trois modules importants dans mes données	23
4. Conclusions de l'analyse bio-informatique	31
5. Caractérisation fonctionnelle des lncRNAs dans des lymphocytes T CD8 ⁺ humains.	32
5.1. Electroporation de siRNA	32
5.2. Internalisation de GapmeR	35
5.3. Impact de l'inhibition de LINC00402 sur les fonctions de lymphocytes T CD8	41
Discussion	43
1. Préambule	43
2. Est-ce que nous avons le matériel adéquat pour répondre à nos objectifs ?	44
3. Qu'est-ce que l'analyse des réseaux de co-expression a apporté de plus que l'analyse de la littérature ?	44
4. Peut-on déduire des fonctions biologiques des 3 modules de gènes co-exprimés étudiés dans ce mémoire ?	45
5. Quelles sont les limitations de l'analyse des réseaux de co-expression par WGCNA?	46

6. Pourquoi les lymphocytes T CD8 ⁺ sont-ils si difficiles à transfecter par siRNA ?	48
7. Comment interpréter l'absence d'inhibition de CD11a à la surface des lymphocytes T CD8 ⁺ traités avec les GapmeR anti-CD11a ?	48
8. Quel est l'impact de l'inhibition de LINC00402 dans des lymphocytes T CD8 ⁺ humains ?	49
9. Est-ce que ma méthode bioinformatique est applicable à d'autres données ?	50
10. Perspectives	52
Matériels et Méthodes	53
1. Analyse bio-informatique	53
1.1. Identification des gènes différentiellement exprimés	53
1.2. Construction des réseaux de co-expressions.	54
1.3. Analyses d'enrichissement fonctionnel	55
2. Cultures cellulaires et milieux	55
3. Décongélation	55
4. Isolement des lymphocytes T CD8 ⁺ du sang de donneurs atteint d'hémochromatose	56
5. Marquages et cytométrie en flux	57
6. Stimulation de lymphocytes T CD8	58
7. Nucléofection de siRNA	58
8. Traitement aux GapmeR	59
9. RT-qPCR	60
10. Test de sécrétion de cytokines	61
11. Test de prolifération	62
Bibliographie	64
Annexes	69
Annexe 1 : Paramètres de l'algorithme de construction des modules d'expression génique	69

Introduction

1. Une brève histoire de l'épuisement des lymphocytes intra-tumoraux

L'interaction entre système immunitaire et cancer est l'objet de nombreuses études depuis des décennies. Ces études ont permis d'identifier les acteurs principaux de la lutte anti-cancéreuse menée par le système immunitaire. Parmi ces acteurs, les lymphocytes T CD8⁺ nous protègent en induisant la lyse des cellules cancéreuses (Coulie et al., 2014). Cette capacité a permis le développement de l'immunothérapie, une approche visant à induire une réponse cytolytique des lymphocytes T contre les cellules tumorales. Cette méthode n'est efficace que dans une fraction des patients (Waldman et al., 2020). Une meilleure compréhension des mécanismes de dysfonctionnement des lymphocytes T CD8⁺ dans leur action anti-tumorale est donc considéré comme une stratégie pertinente pour améliorer les traitements immunothérapeutiques.

De nombreux cancers échappent au contrôle du système immunitaire (Hanahan et Weinberg, 2011). Cette évasion immunitaire est permise, entre autres, par une diminution des fonctions effectrices des lymphocytes T CD8⁺ anti-tumoraux. (Hanahan et Weinberg, 2011, Thommen et al., 2018). Cette dysfonction est associée au concept "d'épuisement", initialement découvert dans un modèle murin d'infection chronique au siècle passé : le modèle LCMV. Dans cette introduction nous verrons comment les caractéristiques clés de l'épuisement ont été découvertes dans le modèle LCMV et comment étudier spécifiquement les lymphocytes épuisés. Nous verrons ensuite le potentiel de l'étude des longs ARN non-codants (lncRNAs) pour mieux comprendre l'épuisement des lymphocytes T CD8⁺.

1.1. Découverte de l'épuisement : Le modèle LCMV

En 1936, Traub et son équipe infectent des souris avec le virus de la chorioméningite lymphocytaire ou LCMV. Ils découvrent alors que les souris qui survivent à l'infection développent une immunité face au virus malgré sa persistance dans l'organisme (Hotchin JE and Cinitis M, 1958; Traub 1936). Cette apparente contradiction a incité Rafi Ahmed et son équipe à étudier cette persistance chronique et à isoler le variant viral responsable, le *clone 13*. Les lymphocytes des souris chroniquement infectées avec le clone 13 présentent une production réduite d'interféron gamma (IFN γ) et une capacité cytotoxique moindre (Gallimore et al., 1998 ; Ahmed et al., 1984). Ces découvertes ont conduit à l'utilisation du terme « épuisé » pour qualifier les lymphocytes aux fonctions diminuées lors d'une infection chronique.

Les recherches menées avec le modèle LCMV ont permis de mieux caractériser le phénomène d'épuisement. Des études omiques suggèrent que le transcriptome et l'épigénome des lymphocytes T épuisés sont distincts des lymphocytes T CD8⁺ effecteurs ou mémoire classiques. (ATAC-seq :Scott-Browne et al., 2016 ; WGCNA Doering et al., 2012). John Wherry et son équipe ont démontré en 2003 que l'épuisement est un phénomène progressif qui s'aggrave avec la durée de l'infection et avec la charge virale. Ils identifient la diminution successive de production d'IL-2, d'IFN γ et de TNF α . Malgré cette dysfonction, les lymphocytes T CD8⁺ épuisés possèdent encore des fonctions effectrices car ils sont toujours nécessaires au contrôle de l'infection virale (Paley et al., 2012). Une autre caractéristique clé des lymphocytes T épuisés suite à l'infection chronique du LCMV est leur expression d'une panoplie de récepteurs inhibiteurs (PD-1, CTLA4, LAG3, ...) et du facteur de transcription TOX (*Thymocyte selection-associated high mobility group box protein*). (Wherry et al., 2007 ; Khan et al. 2019).

Comme vu dans ce chapitre, les études qui décrivent l'épuisement et le développement des lymphocytes T épuisés proviennent du modèle d'infection chronique par le clone 13 du LCMV. Les parallèles avec une situation cancéreuse se sont légitimement posés : que ce soit dans l'infection chronique ou dans le cancer, l'antigène persiste. Mais retrouve-t-on également des lymphocytes T épuisés dans les tumeurs humaines ?

1.2. Existence de l'épuisement dans les tumeurs humaines et murines.

Dans les tumeurs murines, les lymphocytes T intra-tumoraux (TILs) acquièrent un phénotype similaire aux lymphocytes T épuisés étudiés dans le modèle LCMV. En effet, ils co-expriment les récepteurs inhibiteurs typiques des lymphocytes T anti-viraux épuisés : PD-1, LAG-3 et TIM-3. Ils présentent aussi un défaut de production d'IFN γ et de TNF α et expriment fortement TOX. (Khan et al., 2019, Scott et al., 2019).

Au sein des tumeurs humaines, on retrouve des TILs qui partagent de nombreuses caractéristiques avec les TILs murins et les lymphocytes T épuisés du modèle LCMV : l'expression de TOX et de PD-1, la co-expression de multiples récepteurs inhibiteurs et, évidemment, un défaut de production d'IFN γ et de TNF α (Thommen et al., 2015). Toutes ces similarités soutiennent la pertinence du modèle LCMV pour l'étude du dysfonctionnement des lymphocytes T dans les tumeurs humaines.

De plus, certaines discordances ont été soulevées. Par exemple, le *knock-out* de TOX améliore les fonctions des lymphocytes T épuisés dans le modèle LCMV mais ce n'est pas le cas dans un modèle d'hépatocarcinome spontané chez la souris (Scott et al., 2019). Ces discordances suggèrent que de plus amples études sont nécessaires pour clarifier le rôle de TOX dans les TILs et pour découvrir d'autres acteurs moléculaires impliqués dans l'épuisement des lymphocytes T CD8⁺ anti-tumoraux.

Malgré des caractéristiques communes, le dysfonctionnement des TILs CD8⁺ humains n'est pas une copie conforme de l'épuisement des lymphocytes T du modèle LCMV. Il a été notamment montré que les TILs épuisés de NSCLC (*Non-small cell lung carcinoma*) présentent une glycolyse et une absorption de lipide accrue, deux altérations métaboliques qui ne sont pas retrouvées dans les lymphocytes T épuisés du modèle LCMV. De plus, les signatures génétiques décrites à partir des lymphocytes T épuisés du modèle LCMV ne sont pas enrichies dans les TILs épuisés de NSCLC, ce qui suggère que ces deux populations présentent des états de dysfonctionnement distincts (Thommen et al., 2018). De nombreuses différences sont donc à anticiper entre les

lymphocytes T épuisés du modèle murin LCMV et les TILs humains. Ces différences soulignent l'importance d'étudier l'épuisement des TILs chez l'homme.

1.3. Identifier les lymphocytes T épuisés grâce au facteur de transcription TOX

Pour identifier les lymphocytes T épuisés, un marqueur fiable et spécifique de l'épuisement est nécessaire. Une approche commune est d'utiliser l'expression des récepteurs inhibiteurs typiques des lymphocytes épuisés (PD-1, CTLA4, LAG3,...). Malheureusement, de tels récepteurs inhibiteurs sont également exprimés à la surface de lymphocytes T récemment activés (Wherry et al., 2007), ce qui compromet l'utilisation de l'expression des récepteurs inhibiteurs comme marqueur spécifique des lymphocytes épuisés.

Le facteur de transcription TOX nous permet de mieux identifier les lymphocytes T épuisés. Suite au contact entre le récepteur des lymphocytes T (TCR) et un antigène tumoral, plusieurs cascades de signalisation sont déclenchées. Durant ce phénomène, la voie NFAT est activée et le facteur de transcription NFAT (*Nuclear Factor for Activated T-cells*) est internalisé dans le noyau des lymphocytes T et déclenche l'expression de TOX. Les lymphocytes T activés expriment donc TOX. Dans le cas d'une infection persistante, les lymphocytes T CD8⁺ activés présentent des niveaux élevés de TOX supérieurs à ceux observés chez les lymphocytes T CD8⁺ activés circulant dans le sang (Données non publiées du laboratoire d'accueil). L'expression de TOX est associée à une augmentation de l'expression du récepteur inhibiteur PD-1 et à une moindre production d'IFN γ et de TNF α , deux caractéristiques clés de l'épuisement. Elle empêche également l'installation des programmes transcriptionnels typiques des lymphocytes T effecteurs et mémoires et stimule la différenciation des lymphocytes T vers l'épuisement (Khan et al., 2019). Ces caractéristiques suggèrent que l'expression élevée de TOX pourrait constituer un bon marqueur pour identifier les lymphocytes T épuisés.

Pour tester cette hypothèse, le niveau d'expression de PD-1 et de TOX a été analysé dans des échantillons de lymphocytes T CD8⁺ du sang et de la tumeur de patients par cytométrie en flux (**Figure 1**). Une proportion des TILs CD8⁺ de carcinome ovarien expriment PD-1 à un niveau supérieur à celui que l'on peut retrouver dans les T CD8⁺

sanguins (**Figure 1, A**). On distingue donc trois populations au sein des TILs CD8⁺ : la population PD-1 négative (PD-1^{neg}), la population PD-1 positive intermédiaire (PD-1^{int}) et la population PD-1 élevée (PD-1^{high}). Lorsque l'expression de TOX est évaluée dans les TILs CD8⁺, on constate qu'ils l'expriment également à un niveau supérieur à celui des lymphocytes T CD8⁺ sanguins (**Figure 1, A**). On peut donc, ici aussi, distinguer les populations TOX^{neg}, TOX^{int} et TOX^{high}. Lorsque l'on évalue l'expression de TOX des TILs PD-1^{high} et l'expression de PD-1 des TILs TOX^{high}, on peut observer que les TILs PD-1^{high} et les TILs TOX^{high} constitue une seule et même population (**Figure 1, A**). Conformément à ce qui a été décrit dans la littérature, suite à leur activation *in vitro* à l'aide d'anticorps anti-CD3, les TILs TOX^{high} présentent un défaut de production d'IFN γ (**Figure 1.B**) (Thommen et al., 2018¹). Ces observations ont mené le laboratoire à conclure que les TILs PD-1^{high} TOX^{high} sont des lymphocytes avec un phénotype épuisé. Ce constat nous permet donc d'identifier et d'isoler les TILs épuisés par cytométrie en flux uniquement sur base de leur niveau d'expression de PD-1.

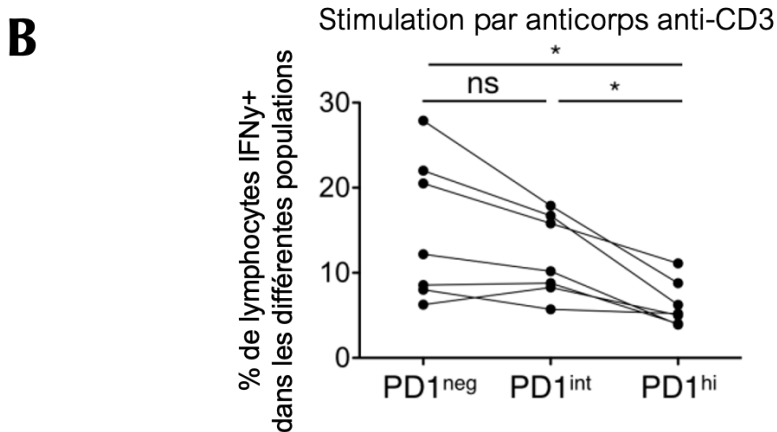
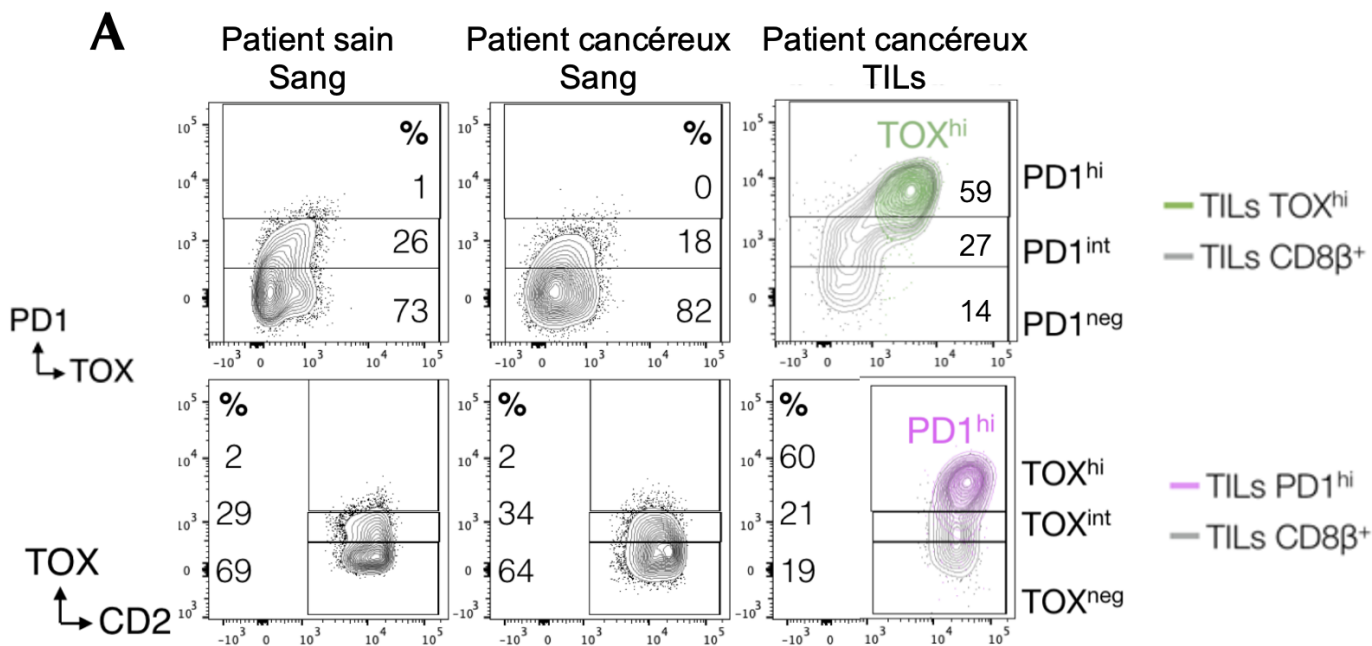


Figure 1 : Les lymphocytes T CD8⁺ PD-1^{high} sont identiques aux lymphocytes T CD8⁺ TOX^{high}.
(A) Stratification des lymphocytes T CD8⁺ en trois populations basées sur l'expression de PD-1 ou de Tox. Les lymphocytes T CD8⁺ PD-1^{high} sont ceux dont le niveau d'expression de PD-1 dépasse le niveau maximal d'expression de PD-1 dans le sang de donneurs non-cancéreux. Les lymphocytes T CD8⁺ TOX^{high} sont ceux dont le niveau d'expression de TOX dépasse le niveau maximal d'expression de TOX dans le sang de donneurs non-cancéreux. (B) Pourcentage de cellules productrices d'IFN γ dans les TILs CD8⁺ PD-1^{neg}, PD-1^{int} ou PD-1^{hi} après stimulation par des anticorps anti-CD3. * = p<0,05. (Tests de Dunn)
Adapté de la thèse de Thibault Hirsch, 2022.

Les conclusions du modèle LCMV et des résultats récents du laboratoire sur les TILs CD8⁺ PD-1^{high} (**Figure 1**) nous permettent donc d'identifier facilement une population enrichie en lymphocytes T épuisés en isolant les TILs CD8⁺ PD-1^{high} par cytométrie en flux. Avec cette technique, nous pouvons étudier spécifiquement le transcriptome des TILs épuisés pour mieux comprendre le phénomène de l'épuisement dans les TILs CD8⁺ humains. Dans cette optique et pour porter un regard innovateur sur l'épuisement, nous nous intéressons à une partie méconnue du transcriptome : les longs ARN non-codants.

2. Les longs ARN non-codants

Les longs ARN non-codants (*long non coding RNA*, lncRNAs) sont des transcrits de plus de 200 paires de bases qui ne codent pas pour une protéine. Selon la base de données NONCODEV5, plus de 95 000 lncRNAs ont été annotés dans le génome humain (Fang, 2018).

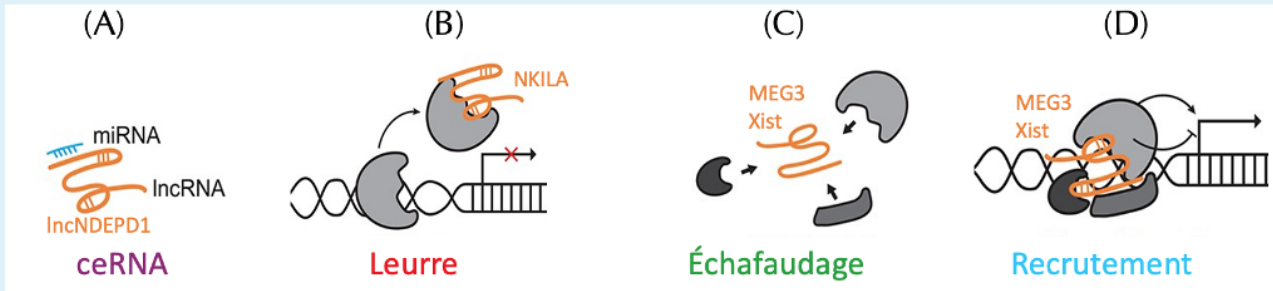
Longtemps considérés comme issus d'un bruit de fond génomique inutile, les lncRNAs sont aujourd'hui de plus en plus considérés comme des acteurs à part entière de la biologie moléculaire. Ils sont capable d'être exprimés, d'interagir avec des brins d'ADN, des protéines et d'autres molécules d'ARN (**Boite 1, page 11**) et sont même sujets à des modifications post-transcriptionnelles capable de moduler leurs fonctions (Chen et al., 2017).

Il est fréquent pour les lncRNAs de réguler l'expression de leurs gènes voisins. La position relative entre un lncRNA et ceux-ci est un facteur clé de leur relation

régulatrice. Cette régulation peut être accomplie par deux mécanismes principaux, qui ne sont pas mutuellement exclusifs : un lncARN peut réguler ses gènes voisins par l'action de son transcrit, et/ou la transcription ou l'épissage du lncARN peut générer une configuration de la chromatine favorable ou défavorable à l'expression des gènes proches, peu importe la séquence du transcrit. Par exemple, le lncRNA CT-GABRA3 chevauche le promoteur du gène GABRA3. Dans une situation cancéreuse, le promoteur de CT-GABRA3 est déméthylé et sa transcription induit le dépôt de marques répressives qui vont mener à l'inhibition épigénétique du promoteur de GABRA3 (Fain et al., 2021). C'est à cause de ce type de mécanismes que les lncRNAs sont souvent distingués selon leur position au sein de leur locus de gènes (**Figure 2**).

Les lncRNAs sont impliqués dans de multiples processus physiologiques et leur expression est particulièrement cellule- ou tissu-spécifique (Atianand et al., 2017). Ces résultats ont suscité un intérêt pour l'étude des lncRNA afin d'en savoir plus sur le fonctionnement du système immunitaire. Les implications connues des lncRNAs dans le système immunitaire sont déjà importantes (Chen et al., 2017). Plusieurs lncRNAs sont notamment connus pour restreindre l'inflammation comme NKILA ou Lethe qui régulent négativement la voie NF-KB (Huang et al., 2018, Chen et al., 2017). On découvre, d'ailleurs, déjà des lncRNAs impliqués directement dans l'épuisement comme lncNDEPD1 qui restreint l'expression de PD-1 en leurrant un de ses microARN (Cheng et al., 2022). Que ce soit dans la maturation, dans l'activation ou même dans l'épuisement des lymphocytes T, les lncRNAs semblent impliqués profondément dans le fonctionnement du système immunitaire. Etant donné le peu d'étude qui s'intéressent aux lncRNAs dans les lymphocytes T intra-tumoraux, on peut supposer que plusieurs lncRNAs encore inconnus sont impliqués dans l'épuisement des lymphocytes T CD8⁺ intra-tumoraux et qu'il ne reste qu'à les découvrir.

Boite 1 : Les mécanismes d'action des lncRNA



(A) Leur fonction la plus connue est celle d'ARN compétitif endogène (ceRNA). En comportant une séquence complémentaire à celle d'un micro-ARN, ils peuvent s'y lier et ainsi inhiber leur activité. C'est notamment le cas de *lncNDEPD1* qui lie un micro-ARN dont la cible est l'ARN messager de PD-1 (Cheng et al., 2022). **(B)** Ils peuvent modifier la fonction d'une protéine en interagissant avec elle. En suivant cette approche, le lncRNA *NKILA* interagit avec le complexe NF- κ B-IKB, ce qui empêche la transduction de la voie NF- κ B au sein des lymphocytes T (Huang et al., 2018). **(C et D)** En interagissant avec plusieurs protéines, un lncRNA peut servir d'échafaudage à un complexe protéique. Le lncRNA *Xist* joue un rôle majeur dans le développement embryonnaire des femelles mammifères en orchestrant la désactivation d'un des deux chromosomes X. Cela est notamment possible car *Xist* participe à la formation d'un complexe protéique capable de déposer des marques épigénétiques répressives sur les histones du chromosome X inactivé. (Kopp et al., 2018) Les lncRNAs peuvent aussi interagir avec des brins d'ADN en formant un triplex triple brin ADN-ADN-ARN (Li et al. 2017) et ainsi recruter diverses protéines sur des régions bien particulières. C'est notamment le cas de *MEG3*, un lncRNA qui recrute le complexe modificateur d'histone PRC2 sur les gènes de la voie de signalisation du TGF-B (Mondal et al., 2015).

Adapté de Park et al. 2022

lincRNA



lncRNA antisens (asRNA)



lncRNA intronique



- Gènes codant une protéine
- Gènes codant un lncRNA

Figure 2 : Localisations génomiques des lncRNA. Les longs ARN non-codants intergéniques (lincRNA) sont positionnés entre deux gènes codants pour des protéines. Les lncRNAs antisens sont exprimés par la transcription antisens (ou divergente) d'un promoteur de gène codant. Néanmoins cette appellation est également utilisée pour décrire les lncRNAs codés par le brin antisens d'un gène codant. Enfin, les longs ARN non-codants introniques sont transcrits à partir d'introns de gènes codants.

Adapté de Atianand et al., 2017.

3. Challenges et objectifs

Les lncRNAs présentent plusieurs particularités qui compliquent leur analyse. Premièrement, l'expression moyenne des lncRNAs est plus faible que celle des gènes codants (Liu et al., 2020 ; Cabili et al., 2015). Deuxièmement, seuls 40 à 50% des lncRNAs possèdent une queue de polyA (**Figure 3**) (Lorenzi et al., 2021). La majorité des kits de purification d'ARN se basent sur la présence de la queue d'adénosine à l'extrémité 3' des molécules d'ARN et sont, par conséquent, incapables de purifier une partie importante des lncRNA.

Troisièmement, près de 70% des lncRNAs ne peuvent pas être ciblés par CRISPR à cause de leur trop grande proximité avec d'autres gènes (Goyal et al., 2017). Il faut donc compter sur d'autres méthodes pour modifier leur expression *in vitro*.

Statut de polyAdénylation des transcrits identifiés

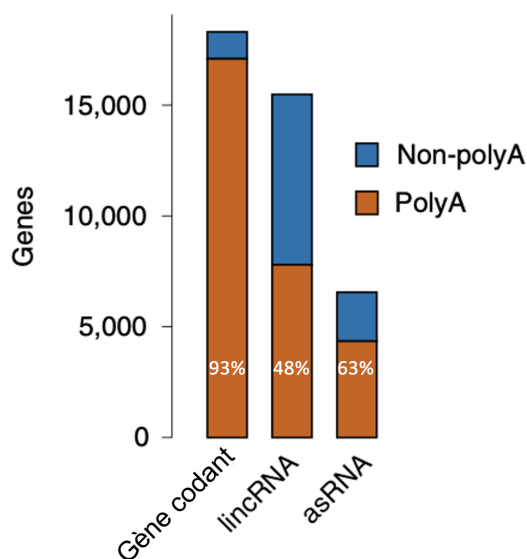


Figure 3 : Une grande partie des longs ARN non-codants ne possèdent pas de queue de polyA. Le nombre de transcrits polyadénylés et non-polyadénylés au sein de 3 biotypes a été déterminé en comparant les données de séquençage obtenues avec une purification d'ARN classique et une purification d'ARN qui conserve les ARN sans queue de polyA. Le pourcentage de transcrits polyadénylé est indiqué en blanc. *lincRNA* : long ARN non-codant intergénique, *asRNA* : long ARN non-codant antisens. Adapté de Lorenzi et al., 2021

Enfin, les analyses d'enrichissement fonctionnels des lncRNAs possèdent une limitation importante : ces analyses se basent sur des données empiriques pour déterminer les familles de gènes liés à différents phénotypes. Or, les longs ARN non-codants sont rarement détectés dans la littérature et sont peu conservés dans d'autres espèces (Quinn et al., 2016). Cela implique que les familles de gènes utilisées dans les analyses contiennent rarement des lncRNA. Nous devons donc trouver des moyens pour étudier l'implication des lncRNAs dans des familles de gènes liés à l'épuisement des lymphocytes (Voir *Analyse des réseaux de co-expression géniques*, page 19).

L'étude des lymphocytes, quant à elle, possède aussi son lot de challenges. Les lymphocytes T humains sont connus pour être "difficiles à transférer" (Yin et al., 2006). Ils sont résistants aux agents de transfections (Yin et al., 2006) et disposent d'une machinerie d'interférence ARN peu efficace (Oberdoerffer et al., 2005, Mantei et al., 2008).

Toutes ces caractéristiques compliquent sérieusement l'identification et la caractérisation fonctionnelle des lncRNAs au sein des lymphocytes T humains.

OBJECTIFS DU MÉMOIRE

Comme nous venons de le voir, les lncRNAs sont impliqués dans de nombreux processus physiologiques et pathologiques, et sont particulièrement actifs dans les fonctions des lymphocytes T. Il semble donc pertinent de rechercher d'éventuels lncRNAs impliqués dans l'épuisement des lymphocytes T intra-tumoraux. De plus, aucune étude à grande échelle des lncRNAs dépourvus de queue de polyA n'a été réalisée jusqu'à présent dans les lymphocytes T intra-tumoraux. Compte tenu de l'existence potentielle de tels lncRNAs encore inconnus et des challenges posés par leur étude dans des lymphocytes T primaires, l'objectif de mon mémoire est double :

- 1) Mettre au point une méthode **bioinformatique** capable d'identifier des lncRNAs liés à l'épuisement des lymphocytes T CD8⁺ intra-tumoraux.
- 2) Développer une méthode **biocellulaire** afin d'interroger le rôle de ces lncRNAs dans la fonction des lymphocytes T CD8⁺ intra-tumoraux.

4. Stratégie mise en place pour accomplir nos objectifs

Pour développer une méthode bioinformatique capable d'identifier des lncRNAs impliqués dans l'épuisement, un jeu de données déjà disponible au sein du laboratoire a été utilisé (**Figure 4**). Ces données sont des transcriptomes de lymphocytes T CD8⁺ d'ascites de patients cancéreux et du sang de donneurs atteint d'hémochromatose. L'ascite carcinomateuse est un liquide qui s'accumule dans la cavité abdominale de patients atteints de certains cancers. Elle se forme généralement à la suite d'une situation inflammatoire dans l'abdomen, causée par la présence de la tumeur. On y retrouve des cellules du système immunitaire, et notamment des lymphocytes T CD8⁺. Les transcriptomes de ces lymphocytes et de lymphocytes du sang de donneurs non-cancéreux ont été comparés.

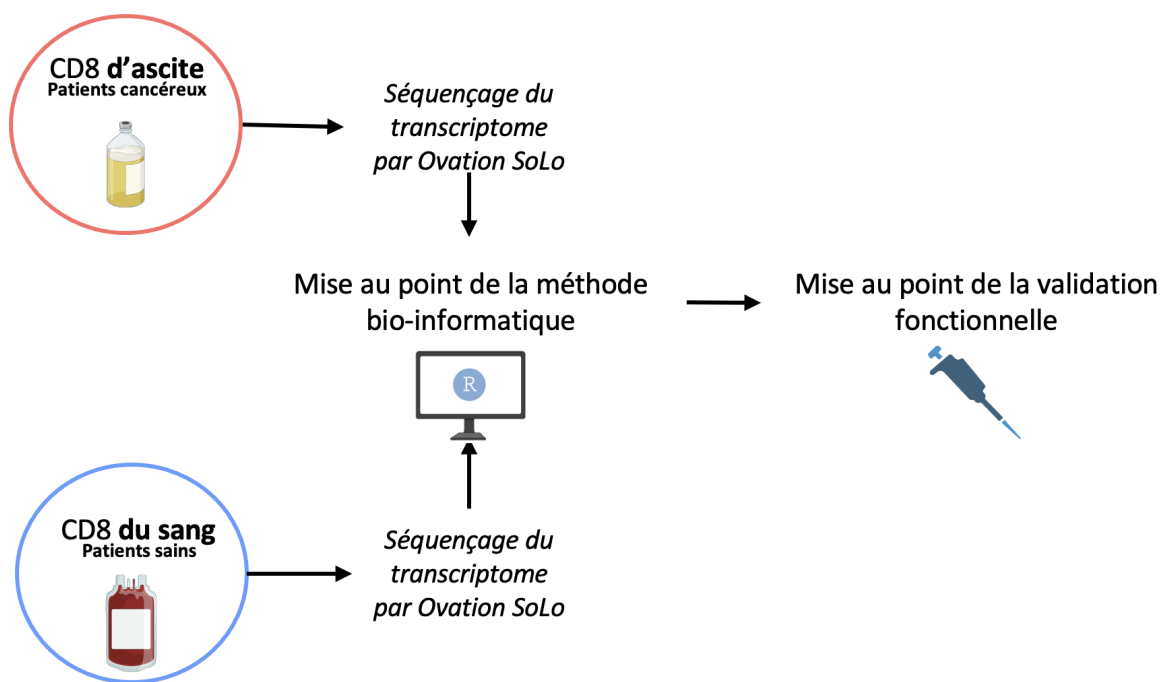


Figure 4 : Objectifs et stratégie. Les lymphocytes T CD8⁺ isolés à partir d'ascite carcinomateuses et du sang de donneurs non-cancéreux ont été séquencés avec le kit Ovation SoLo. Ce kit permet de inclure l'ensemble des lncRNAs dans notre analyse. Ces données seront utilisées pour réaliser les deux objectifs de ce mémoire. Réalisé avec BioRender.com

Ces données présentent plusieurs avantages pour mettre au point une méthode bioinformatique. Premièrement, elles ont été générées grâce au kit Ovation SoLo, ce qui nous permet de détecter les lncRNAs sans queue de polyA et donc de travailler avec l'ensemble des transcrits qui nous intéressent. Ensuite, il existe des lymphocytes T CD8⁺ dysfonctionnels dans les ascites. (Demotte et al., 2010). Certaines études parlent de

l'existence de lymphocytes T épuisés (Nakano et al., 2018). Mon travail de mise au point à partir de ces données nous permet donc également de rechercher les signatures transcriptomiques de lymphocytes T épuisés dans les ascites tumorales.

Ces données n'ont pas été générées pour convenir à mon projet de mémoire et présentent donc certains désavantages. Des lymphocytes T naïfs sont présents dans ces échantillons. Les lymphocytes T naïfs sont des lymphocytes T qui n'ont pas encore rencontré leur antigène. Leur présence risque de masquer les signatures transcriptomiques des lymphocytes T dysfonctionnels.

Résultats

1. Aperçu général de mes données

Un graphique d'analyse en composantes principales nous permet d'avoir une idée du comportement global de l'expression génique des lymphocytes T CD8⁺ dans nos échantillons (**Figure 5**). Les deux conditions forment des populations clairement délimitées et ne se chevauchent pas. Le graphique ne représente que 63% de la variabilité totale des échantillons, ce qui témoigne d'une certaine hétérogénéité dans nos données.

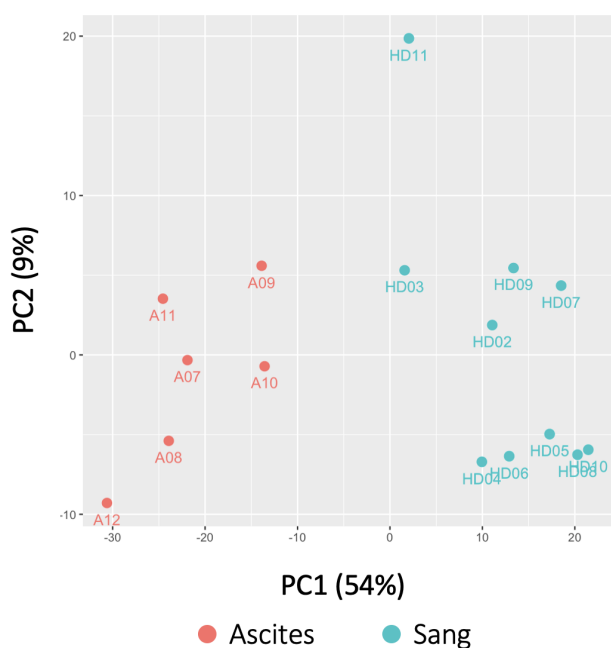


Figure 5 : Graphique d'analyse en composantes principales des données transcriptomiques de mes échantillons. (Ascite), Lymphocytes T CD8⁺ isolés d'ascites tumorales. (Sang), Lymphocytes T CD8⁺ isolés du sang de donneurs non-cancéreux.

2. Définition des lncRNAs différentiellement exprimés

Nous avons identifié 486 gènes différentiellement exprimés dans nos données. Parmi eux, 94 gènes codent des lncRNAs (**Figure 6**). Nous avons mené plusieurs approches pour réduire la taille de cette liste. Premièrement, une analyse de la littérature de ces lncRNAs a été conduite afin d'identifier les transcrits qui pourraient présenter des fonctions dans l'épuisement des lymphocytes T CD8⁺. Il est important de noter que notre étude de la littérature ne couvre pas tout ce qui a été publié sur ces lncRNAs et qu'il existe de nombreux autres travaux qui auraient pu être pris en compte. Les lncRNAs peuvent réguler l'expression de leurs gènes voisins (Introduction, pages 9 et 10). Une analyse de la littérature doit donc aussi porter sur ces gènes. Au terme de cette analyse (**Tableau 1, page 18**), 11 lncRNAs se distinguent par leur fonctions potentielles dans l'épuisement

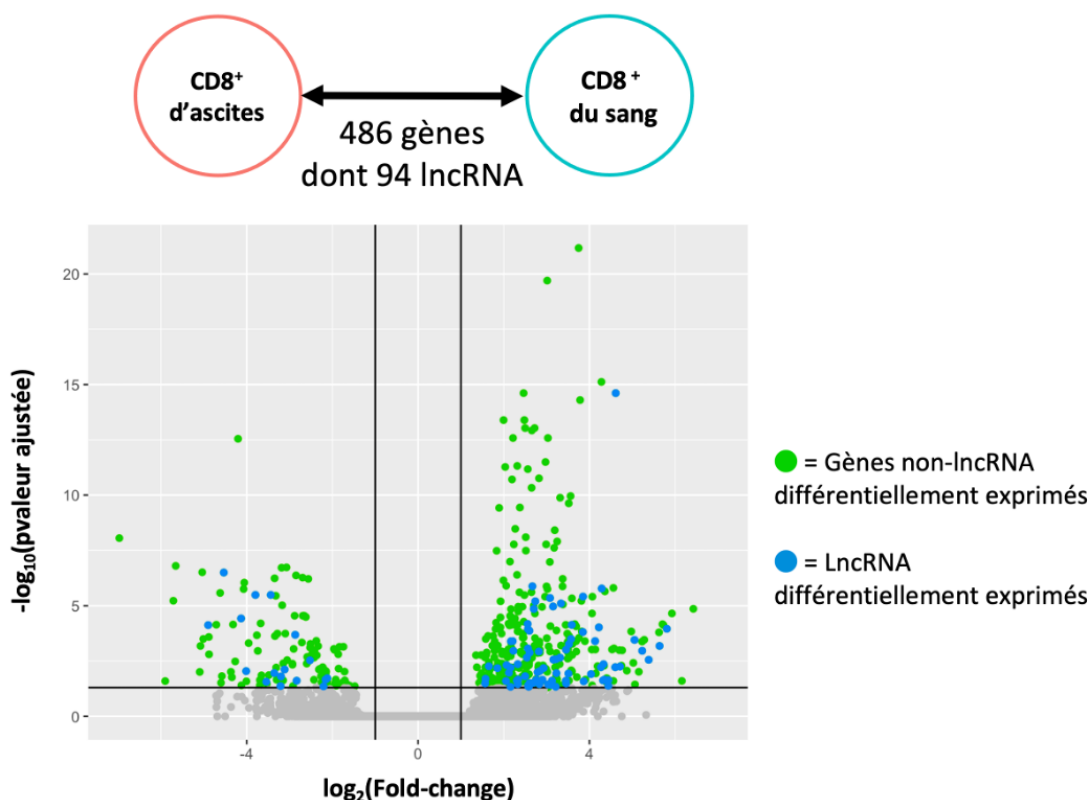


Figure 6 : Volcano Plot des gènes différentiellement exprimés entre nos 2 conditions. Le Fold-change correspond à l'intensité de la différence d'expression des gènes entre les deux conditions. La p valeur ajustée correspond à la p valeur de l'expression différentielle calculée après avoir éliminé les gènes parasites et appliqué la correction de Benjamini-Hochberg. Chaque point représente un gène. Les barres noires définissent les seuils utilisés pour définir les gènes différentiellement exprimés. Les barres verticales correspondent à des Fold-change de 0,5 et 2. La barre horizontale correspond à une p valeur de 0,05.

Nom	Log2 (Fc)*	Littérature	Références
<i>SATB1-AS1</i>	+2,7	<p>Son expression est corrélée à celle de SATB1 dans les cancers colorectaux</p> <p>Antisens de SATB1 SATB1 est impliqué dans l'épuisement et dans la maturation des lymphocytes T. Il forme un complexe qui restreint l'expression de PD-1. Cette inhibition est inhibée par le TGF-β.</p>	<p>Shademan et al., 2019</p> <p>Nixon et li., 2017</p>
<i>LINC00402</i>	+4,6	<p>Est impliqué dans l'activation des lymphocytes T en réponse à un stimulus allogénique. Il stimule l'activité de ERK1 et ERK2 ainsi que l'expression d'IL-2 et EGR-1</p> <p>Est différentiellement exprimé dans une analyse comparant des TILs CD8⁺ PD-1^{high} vs PD1^{int} de cancer bronchique non à petite cellules</p>	<p>Peltier et al., 2021</p> <p>Thommen et al., 2018</p>
<i>SIRPα-AS1</i>	+3,2	<p>Antisens de SIRPα SIRPα est un récepteur transmembranaire. Il interagit avec CD47 à la surface des cellules présentatrices d'antigène et stimule l'activation des lymphocytes T. (CD47 participe à la costimulation et à l'adhésion cellulaire des lymphocytes T)</p> <p>SIRPα est aussi un des 3 gènes dont l'expression corrèle le plus avec l'expression de PD-1 dans des échantillons de carcinome séreux de haut grade dans l'ovaire ($r = 0.90$).</p>	<p>Dehmani et al., 2021</p> <p>Wang et al., 2022</p>
<i>A1BG-AS1</i>	+2,8	<p>Voisin proche de BACH2 BACH2 est un facteur de transcription exprimé par les TILs épuisés humains. Il participerait à l'épuisement par son inhibition de BATF.</p>	Cunningham et al., 2022
<i>TMC3-AS1</i>	+3,5	<p>Son expression est induite suite au traitement de macrophages avec le LPS et inhibe l'expression d'IL-10, une cytokine anti-inflammatoire.</p>	Ye et al., 2020
<i>SRP14-DT</i>	+2,5	<p>Voisin proche de GCN2 GCN2 est essentiel pour la survie et la réponse des lymphocytes T CD8⁺ dans un modèle murin de gliome malin.</p> <p>GCN2 est impliqué dans l'immunosuppression du microenvironnement tumoral.</p> <p>GCN2 participe aussi à la répression de l'inflammasome en réponse à la privation d'acides aminés dans des cellules présentatrices d'antigènes et endothéliales intestinales.</p>	<p>Rashidi et al., 2019</p> <p>Jo Halaby et al., 2019</p> <p>Ravindran et al., 2016</p>

PRKCQ-AS1	+2,1	<p>Son expression est corrélée à celle de PRKCQ dans les cancers colorectaux</p> <p>Antisens de PRKCQ PRKCQ code la kinase PKCθ, une kinase essentielle dans la fonction et la différenciation des lymphocytes. Une perte de PKCθ dans des lymphocytes induit notamment une production d'IL-2 et une prolifération <i>in vitro</i>.</p>	<p>Shademan et al., 2019</p> <p>Marsland et Kopf, 2008</p>
LINC00649	+2,2	<p>Son expression stimule l'activation et la prolifération des lymphocytes T CD4 dans la maladie inflammatoire chronique de l'intestin (IBD) en leurrant miR-125a, un microARN inhibiteur de l'expression d'IL-23R.</p>	Nie et al., 2020
CDKN2B-AS1	-2,1	<p>ANRIL, ou CDKN2B-AS, interagit avec les complexes modificateurs d'histones PRC1 et PRC2 et mènent à l'inhibition épigénétique d'une grande série de gènes. Il régule notamment l'activation du complexe NF-κB dans des lymphocytes T leucémiques en formant un triplex ANRIL/EZH2/p65.</p> <p>Il a été extensivement étudié dans plusieurs pathologies (Cancers, athérosclérose, diabète,...)</p>	<p>Song et al., 2018</p> <p>Kong et al., 2018</p>
CCL3-AS1	-3,8	<p>Antisens de CCL3 CCL3 (aussi appelé MIP1α) est une chémokine impliquée dans l'inflammation en stimulant dans le recrutement et l'activation de macrophages, monocytes et neutrophiles.</p>	Ntanasis-Stathopoulos et al., 2020
MIR155HG	-2,2	<p>MIR155HG (miR155 Host Gene) est le gène hôte du lncRNA homonyme et du microARN miR155.</p> <p>Le lncRNA MIR155HG est exprimé en réponse à la signalisation des voies NF-κB et AP-1, deux voies majeures dans l'activation des lymphocytes T</p> <p>- Est différentiellement exprimé dans une analyse comparant des TILs CD8⁺ PD-1^{high} vs PD1^{int} de cancer bronchique non à petite cellules</p>	<p>Tili et al., 2007</p> <p>Thommen et al., 2018</p>

● = Sur-exprimé dans le sang ● = Sur-exprimé dans les ascites

Tableau 1 : Les 11 lncRNAs potentiellement impliqués dans l'épuisement des lymphocytes T CD8⁺. L'étude de la littérature de nos 96 lncRNAs différentiellement exprimés a révélé 11 lncRNAs potentiellement impliqués dans la réponse immunitaire des lymphocytes T CD8⁺, que ce soit directement ou par la régulation potentielle d'un gène voisin. **Log₂(Fc)*** = Le logarithme en base 2 du fold-change sang-ascite indique le sens de la différence d'expression.

La suite logique de cette analyse est la caractérisation fonctionnelle de ces 11 lncRNAs en modulant leur expression *in vitro*. Dans cette optique, s'il fallait choisir un trio de lncRNAs par lesquels commencer nous choisirions *SATB1-AS1*, *LINC00402* et *LINC00649* sur base de la littérature et de leur profil d'expression. En effet, ils sont inhibés dans les lymphocytes T CD8⁺ d'ascites tumorales et il est plus facile d'inhiber l'expression d'un lncRNA que de la stimuler.

Nous avons voulu pousser l'analyse bioinformatique plus loin et combiner cette première approche avec l'analyse des gènes codants qui sont co-exprimés à nos lncRNAs d'intérêt. Nous avons donc mis au point une analyse des réseaux de gènes co-exprimés. (*Weighted Gene Co-expression Networks Analysis*, WGCNA) (Langfelder et al., 2008)

3. Analyse des réseaux de co-expression génique

Afin de repérer les lncRNAs qui pourraient être impliqués dans l'épuisement, nous avons examiné les gènes codants co-exprimés avec nos lncRNA. Notre but était d'identifier des réseaux de gènes co-exprimés qui incluent nos lncRNAs d'intérêt et des gènes codants dont la fonction est connue, afin de mieux détecter des lncRNAs impliqués dans le fonctionnement des lymphocytes T CD8⁺.

3.1. Identification des modules de co-expressions

La construction et l'analyse des réseaux de gènes co-exprimés a été menée avec le package WGCNA (*Weighted Gene Co-Expression Network Analysis*, version 1.71) (Langfelder et Horvath, 2008 et 2012). Les profils d'expression des 20 000 gènes présents dans nos données ont été utilisés pour calculer un score de co-expression entre chaque paire de gènes (**Boîte 2 : Qu'est-ce qu'une analyse des réseaux de gènes co-exprimés ?**, page 20). Ce score nous permet de construire un dendrogramme à partir duquel les modules de co-expressions sont délimités. Par souci de clarté, une boîte "Interpréter le dendrogramme de co-expression" est présentée à la page 22.

Au terme de l'analyse par WGCNA, nous avons identifié 47 modules (ou réseaux) de gènes co-exprimés. Ces modules sont visualisés sous la forme d'un "dendrogramme de co-expression" (**Figure 7**) (**Boîte 3 : Interpréter un dendrogramme de co-expression, page 22**). Nos 94 lncRNAs d'intérêts sont répartis dans 12 de ces modules. Les différents modules sont nommés par un code couleur arbitraire. De manière intéressante, la majorité des gènes différentiellement exprimés sont regroupés dans 3 modules : les modules "brun", "vert" et "turquoise". De plus, la majorité des 94 lncRNAs différentiellement exprimés fait aussi partie de ces trois modules (**Figure 7**). Ceux-ci contiennent respectivement 28, 13 et 13 des 94 lncRNAs différentiellement exprimés. L'analyse de ces modules nous permet donc d'étudier les fonctions cellulaires qui diffèrent entre nos deux conditions. De plus, la majorité des lncRNAs prometteurs identifiés par l'analyse de la littérature fait partie de ces 3 modules (**Figure 8**). C'est pourquoi ces trois modules ont été étudiés plus en profondeur dans la suite de ce mémoire.

Boîte 2 : Qu'est-ce qu'une analyse des réseaux de gènes co-exprimés ?

Deux gènes sont dits *co-exprimés* lorsqu'ils présentent un profil d'expression similaire au sein des échantillons étudiés. Les niveaux d'expression de ces deux gènes augmentent et diminuent ensemble dans les échantillons. Une manière simple de quantifier cette co-expression est de calculer la corrélation entre l'expression des deux gènes. L'expression d'un facteur de transcription est souvent corrélée avec celle de ses gènes cibles (Albert, 2005). Cette technique permet donc, par exemple, d'identifier les gènes sous le contrôle d'un facteur de transcription.

Une analyse de co-expression génique revient à calculer un score pour chaque paire de gènes et à regrouper les gènes aux profils d'expression les plus similaires dans des ensembles communément appelés "réseaux" ou "modules" de gènes. Malgré le caractère strictement corrélatif de ces réseaux, leur composition indique souvent des gènes actifs dans le même processus biologique (Albert, 2005, van Dam et al., 2018) et leur analyse peut donc nous renseigner sur la fonction de ces gènes, codant ou non.

Le détail du calcul de ce score de co-expression est décrit dans la section Matériel et Méthode, à la page 54.

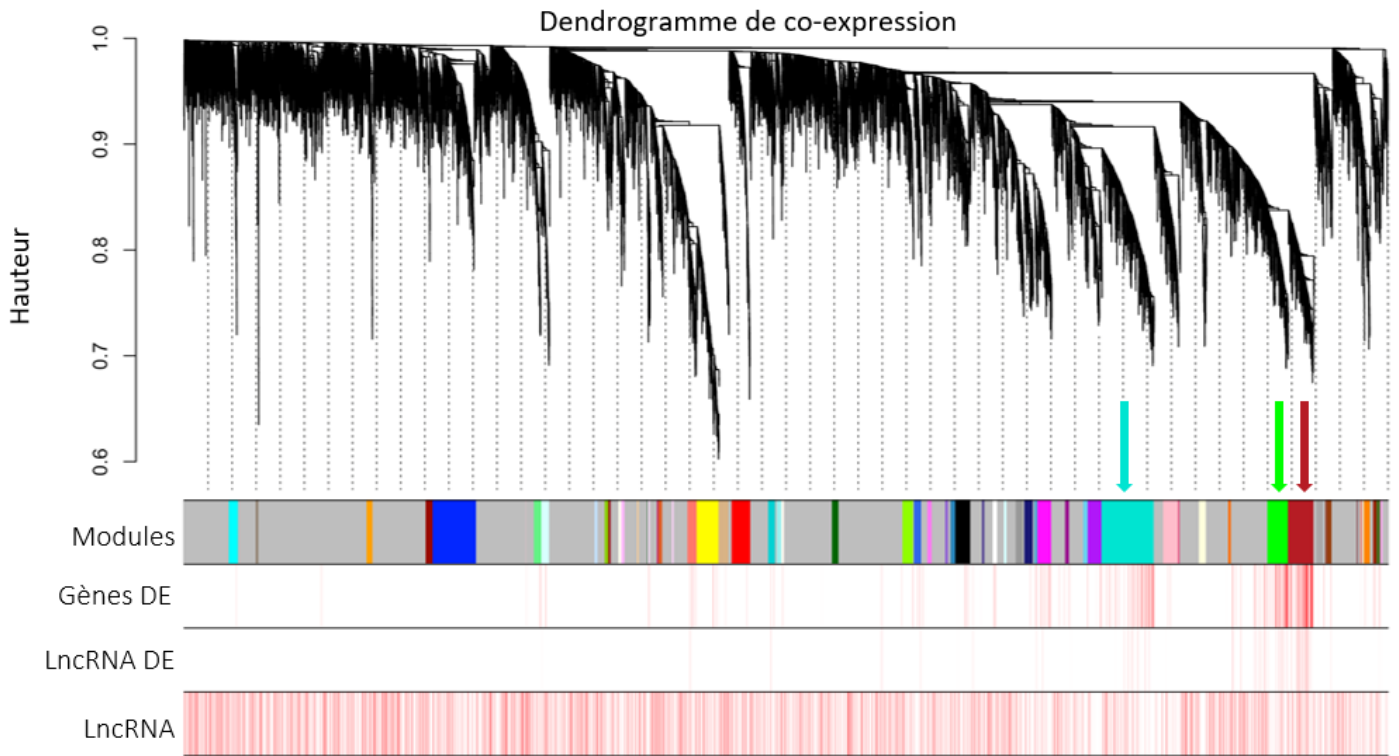
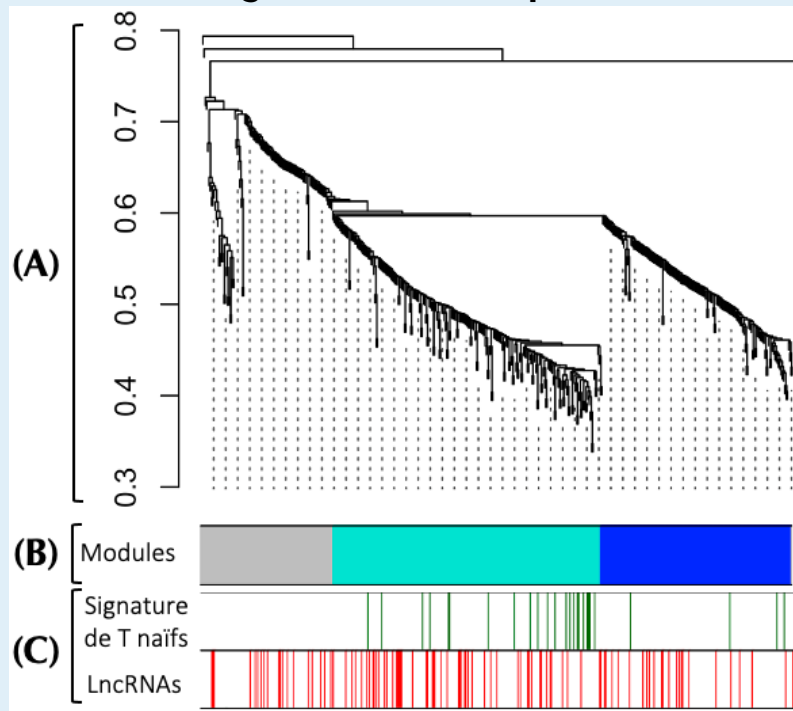


Figure 7 : Construction des réseaux de co-expression génique à partir de nos données. Un dendrogramme des scores de co-expression entre chaque gène nous permet de visualiser les modules de gènes co-exprimés sous forme de branches. Chaque gène est associé ou non à un module par une légende de couleurs. Les trois modules qui incluent la majorité des gènes différentiellement exprimés sont indiqués par des flèches. **Gènes DE** : Gènes différentiellement exprimés. **LncRNA DE** : lncRNAs différentiellement exprimés. **LncRNA** : L'ensemble des lncRNAs identifiés.



Figure 8 : Appartenance des 11 lncRNAs potentiellement impliqués dans l'épuisement des lymphocytes T CD8⁺ aux modules de co-expression de notre analyse. Les couleurs correspondent aux modules ■vert, ■rose, ■brun, ■turquoise et ■bleu royal

Boîte 3 : Interpréter un dendrogramme de co-expression.



(A) Le dendrogramme représente tous les gènes de l’analyse, regroupés selon la similarité de leur expression. Les gènes sont regroupés en cluster de gènes par une barre horizontale. Au plus cette barre est basse, au plus les profils d’expression des gènes regroupés sont similaires. Les modules de gènes co-exprimés prennent la forme de “branches” tombantes.

(B) Un algorithme détermine les limites de ces branches/modules et identifie ainsi les modules qui sont visualisés sous le dendrogramme par différentes couleurs. La couleur grise correspond aux gènes qui n’ont été attribués à aucun module.

(C) Plusieurs légendes de couleurs supplémentaires sont souvent juxtaposées au dendrogramme. Elles permettent d’annoter différents groupes de gènes afin de visualiser leur répartition au sein des différents modules. On peut voir dans cet exemple que les gènes d’une signature de gènes de lymphocytes T naïfs (Szabo et al., 2018) semblent être regroupés dans le module turquoise. Les longs ARN non-codants, quant à eux, semblent être dispersés dans tous les modules.

3.2. Prédiction fonctionnelles de trois modules importants dans mes données

A. Tendance principale et gènes “hubs”

L’*eigengene* ou *gène propre* est un outil essentiel pour synthétiser l’expression d’un module de gènes. Il s’agit de la première composante principale de l’expression des gènes d’un module. Autrement dit, il s’agit d’un vecteur qui décrit au mieux les changements d’expression d’un module dans chaque échantillon. Il permet de décrire

la tendance principale de l'expression du module. Par exemple, les *gènes propres* des modules **■ brun**, **■ vert** et **■ turquoise** présentent globalement une surexpression dans les échantillons de sang de patient sain et une sous-expression dans les échantillons d'ascites. **(Figure 9)** La tendance principale du module **■ turquoise** est systématiquement élevée dans les échantillons sanguins et diminuée dans les échantillons d'ascites, là ou celles des modules **■ brun**, et **■ vert** présentent quelques exceptions.

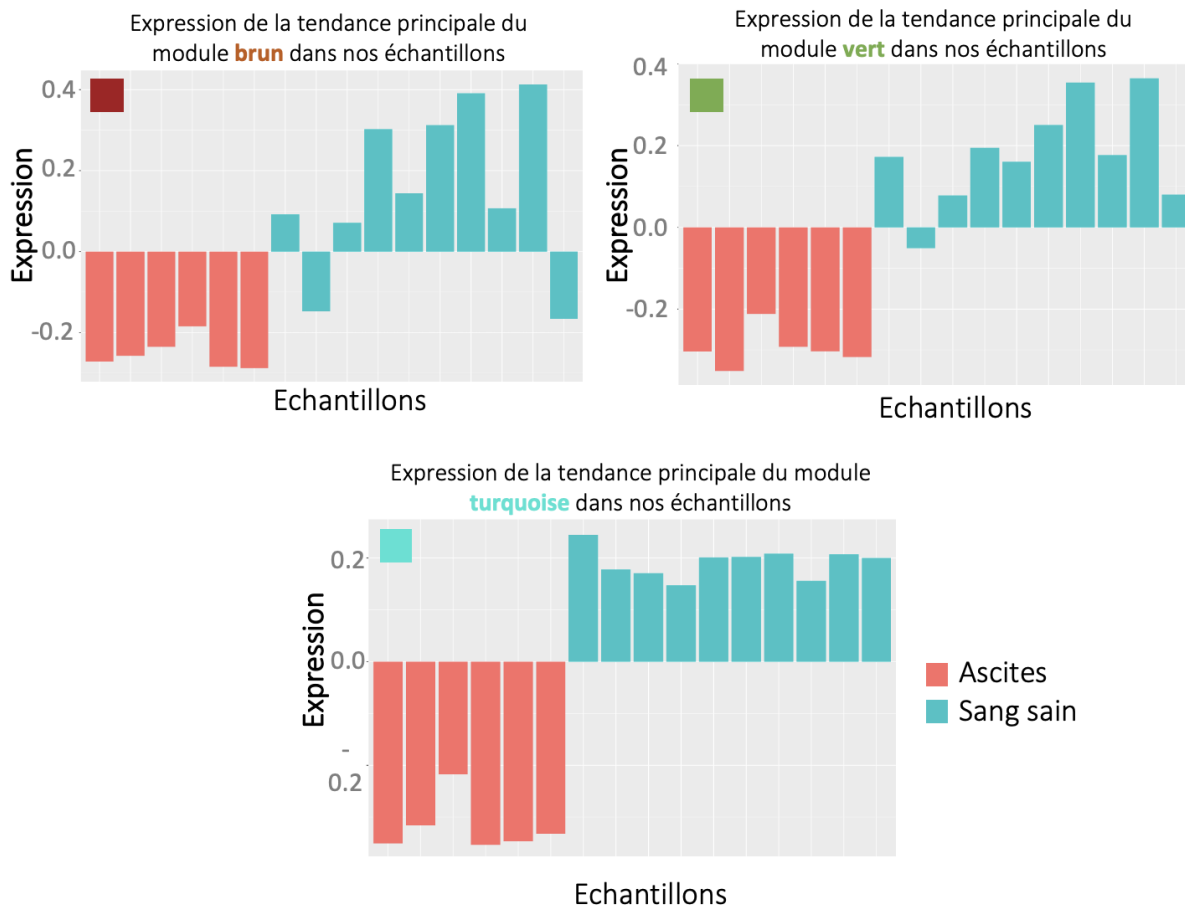


Figure 9 : Tendance principale des 3 modules majeurs de notre analyse. Le profil d'expression de leur gène propre représente la tendance principale d'expression des modules **■ brun**, **■ vert** et **■ turquoise** au sein de nos échantillons.

L'identification des "*gènes hubs*" d'un module de co-expression est un outil utile pour caractériser la fonction biologique commune des gènes d'un module. Il s'agit des gènes qui suivent au plus près la tendance principale du module et qui ont donc plus de chance d'être impliqués dans sa fonction biologique (van Dam et al., 2018). Pour les identifier, un score d'appartenance au module (*Module Membership*) est calculé comme la corrélation entre leur profil d'expression et celui de la tendance principale du module.

Pour les modules ■ brun, ■ vert et ■ turquoise, nous avons identifié les 10 gènes qui ont le plus grand score d'appartenance. Une vue d'ensemble des fonctions connues de ces gènes dans la littérature selon la base de données *GeneCards* a été réalisée (Stelzer et al., 2016) (**Figure 10.B-D**).

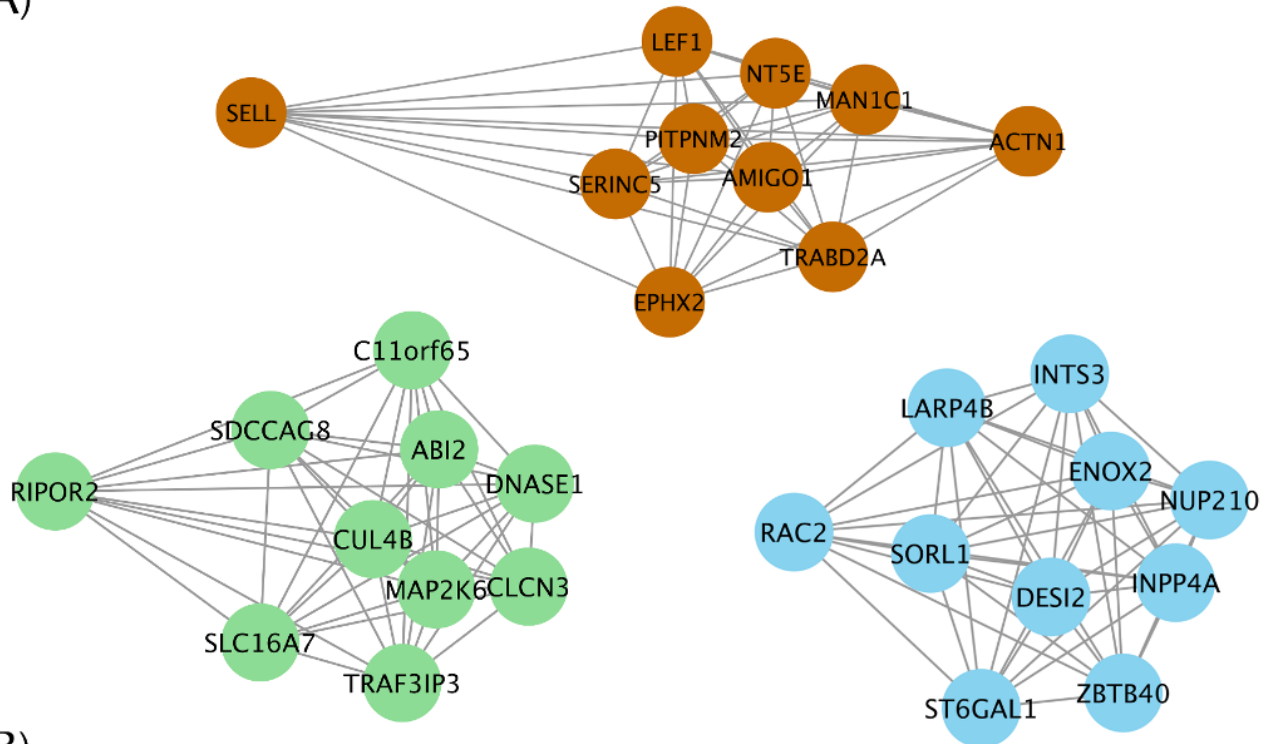
Dans les gènes hubs du module ■ turquoise, *RAC2* et *SORL1* modulent respectivement l'activation du TCR et la signalisation de l'IL-6, une cytokine pro-inflammatoire. Parmi les autres gènes hubs, *INTS3*, *INPP4A* et *ENOX2* sont impliqués dans la croissance cellulaire.

Plusieurs points communs réunissent les gènes hubs du module ■ brun. Premièrement, *LEF1* et *TRABD2A* participent à la voie Wnt (Wikipedia contributors, “*TCF/LEF family*”, 2022 ; *Li et al.*, 2022). Cette voie est notamment impliquée dans la prolifération et la différenciation cellulaire. Ensuite *LEF1* et *CD62L* sont impliqués dans le fonctionnement des lymphocytes T naïf en assurant respectivement la quiescence et le *homing* (migration) de ces lymphocytes aux sites inflammatoires et aux organes lymphoïdes secondaires. La quiescence correspond à est un état stable dans lequel les cellules ne se divisent pas.

Le meilleur gène hub du module ■ vert est *RIPOR2*, un gène codant un inhibiteur d'une RhoA GTPase qui participe au maintien des lymphocytes T naïfs dans un état quiescent en limitant leur prolifération et la sensibilité de leur TCR (Froehlich et al., 2016). *TRAF3IP3*, quant à lui, restreint le métabolisme glycolytique en stimulant l'activité de mTORC1, un acteur majeur de la sortie des lymphocytes T naïfs de leur quiescence suite à leur activation (Chapman et Chi., 2018). Les gènes *SDCCAG8*, *DNASE1*, *MAP2K6* et *ABI2*, quant à eux, présentent des fonctions qui soutiennent la prolifération cellulaire.

Figure 10 : Présentation des gènes hubs des modules ■ brun, ■ vert et ■ turquoise. (A) Représentation en réseau des 10 gènes hubs les plus forts de chaque module. Les réseaux sont construits selon l'algorithme : “Edge-weighted Spring-Embedded Layout” (les gènes sont reliés par des ressorts virtuels qui exercent une force de rappel proportionnellement à l'intensité de la co-expression entre les deux gènes. Plus un gène est fortement co-exprimé à ses voisins, plus il sera repoussé par ceux-ci.) Réalisé avec Cytoscape. (B à D) Aperçu de la littérature des 10 gènes hubs majeurs des modules ■ brun, ■ vert et ■ turquoise (*GeneCards*, Stelzer et al., 2016). Le score d'appartenance à un module correspond à la corrélation entre l'expression du gène hub et la tendance principale du module en question.

(A)



(B)

Les gènes Hubs du module **brun**

Nom	Score d'appartenance au module brun	Fonctions connus
LEF1	0,994	Lymphoid Enhancer Binding Factor 1 Facteur de transcription capable de stimuler l'enhancer alpha du TCR. Il est aussi impliqué dans la voie Wnt.
TRABD2A	0,989	TraB Domain Containing 2A Métalloprotéase membranaire impliquée dans la régulation négative de la voie Wnt. Elle inactive certaines protéines Wnt par clivage.
ACTN1	0,987	Actinin Alpha 1 Alpha-actinine qui se fixe sur l'actine et l'ancre à une variété de structures intracellulaires.
SERINC5	0,987	Serin Incorporator 5 Permet le transport de L-sérine dans le cytosol. Impliqué dans l'immunité innée et dans l'inhibition de l'infection par des lentivirus.
EPHX2	0,987	Epoxide Hydrolase 2 Potentiellement impliquée dans la production des leukotrienes
PITPNM2	0,987	Phosphatidylinositol Transfer Protein Membrane Associated 2 Protéine membranaire qui catalyse le transfert de phosphatidylinositol et de phosphatidylcholine entre les membranes. Lie les ions calcium.
CD62L / SELL	0,985	Selectin L Homing des lymphocytes T aux sites inflammatoires et aux organes lymphoïdes secondaires. Présent sur les lymphocytes naïfs et central mémoire.
CD73 / NTE5	0,983	5'-Nucleotidase Ecto Hydrolyse extracellulaire de nucléotides en nucléosides perméables à la membrane.
MAN1C1	0,983	Mannosidase Alpha Class 1C Member 1 Mannosidase impliqué dans la maturation des oligosaccharides "Asn-linked"
AMIGO1	0,983	Adhesion Molecule With Ig Like Domain 1 Molécule d'adhésion impliquée dans transport du canal potassique voltage-dépendant KCNB1

(C)

Les gènes Hubs du module **vert**

Nom	Score d'appartenance au module vert	Fonction connues
C11orf65	0,989	Chromosome 11 Open Reading Frame 65
RIPOR2/ FAM65B	0,989	Régulation négative de la fission mitochondriale RHO Family Interacting Cell Polarization Regulator 2 Inhibiteur d'une petite GTPase RhoA qui joue de nombreux rôle dans la régulation des myoblastes, la prolifération des lymphocytes T et la polarisation des neutrophiles. Inhibe les réponses des lymphocytes "chémokines-induites" comme la chémotaxie, la polarisation, l'adhésion et la migration cellulaire. Ses deux isoformes majeures sont inhibés dans les lymphocytes T suite à la signalisation du TCR Maintient les lymphocytes T naïfs dans un état quiescent en limitant leur prolifération et la sensibilité de leur TCR.
SDCCAG8	0,986	Protéine associée au centrosome impliquée dans l'organisation du centrosome durant la mitose.
TRAF3IP3	0,986	TRAF3 Interacting Protein 3 - Protéine adaptatrice qui jouent de nombreux rôles dans l'immunité innée et adaptative. - Elle est nécessaire dans le développement des thymocytes murins - Elle restreint le métabolisme glycolytique en facilitant l'interaction entre PP2Ac et mTORC1 dans les lymphocytes T régulateurs. - Elle active le facteur de transcription IRF3 et la production d'interféron en réponse à une infection virale
DNASE1	0,984	Désocytiribonucléase 1 - Endonucléase sérique sécrétée dans plusieurs fluides biologiques (urine et sputum) - Elle bloque la polymérisation de l'actine en se liant à la G-actine
CLCN3	0,983	Chloride Voltage-Gated Channel 3 Antiporteur à chlorure voltage-dépendant qui stimule l'échange de chlorure en échange de protons au niveau de la membrane plasmique et des membranes des vésicules endo-lysosomales.
MAP2K6	0,981	Mitogen-Activated Protein Kinase Kinase 6 - MAP kinase impliquée dans de nombreux processus tels que l'arrêt du cycle cellulaire, l'activation de la transcription et l'apoptose. - Elle participe à la réponse au TNF et à l'IL-12 - Elle est cruciale pour la signalisation de l'IL-6 et l'IFN γ
CUL4B	0,980	Cullin 4B Forme un complexe E3 ubiquitine ligase qui participe à la dégradation de protéines dans de nombreux processus. - Elle régule la voie de signalisation de mTORC1 en interagissant avec MLST8 et DDB1.
ABI2	0,979	Abl interactor 2 - Régule la dynamique du cytosquelette dans la motilité et l'adhésion cellulaire.
SLC16A7	0,978	Solute Carrier Family 16 Member 7 Transporteur de monocarboxylates à chaîne courte tels que le L-lactate, les corps cétoniques et le pyruvate à travers la membrane plasmique.

(D)

Les gènes Hubs du module **turquoise**

Nom	Score d'appartenance au module turquoise	Nom complet et fonctions connues
RAC2	0,988	Rac Family Small GTPase 2 Rho-GTPase dont la perte dans des lymphocytes T cause une baisse de prolifération en réponse à l'activation du TCR. Régule la sécrétion, la phagocytose et la polarisation cellulaire.
NUP210	0,984	Nucleoporin 210 Glycoprotéine majeure de la structure du pore nucléaire
INTS3	0,980	Integrator Complex Subunit 3 Stabilité génomique et régulation du processing de petits ARN nucléaires
ST6GAL1	0,980	ST6 Beta-Galactoside Alpha-2,6-Sialyltransferase 1 Glycosyltransférase qui transfère de l'acide sialique vers des substrats accepteurs contenant du galactose.
DESI2	-0,977	Desumoylating Isopeptidase 2 Déubiquitynilase
SORL1	0,976	Sortilin Related Receptor 1 Récepteur de tri de plusieurs molécules. Régule négativement la signalisation du récepteur membranaire de l'IL-6 tout en stimulant la signalisation du récepteur soluble de l'IL-6. L'IL-6 étant une cytokine pro-inflammatoire.
INPP4A	0,976	Inositol Polyphosphate-4-Phosphatase Type I A Antagonise la voie PI3K-AKT et inhibe donc le cycle cellulaire et la survie cellulaire.
ENOX2	0,975	Ecto-NOX Disulfide-Thiol Exchanger 2 NADH oxidase, peut-être impliquée dans la croissance cellulaire
LARP4B	0,975	La Ribonucleoprotein 4B Ribonucléoprotéine qui stimule la traduction.
ZBTB40	0,974	Zinc Finger And BTB Domain Containing 40 Protéine nucléaire impliquée dans la réponse aux dommages de l'ADN

B. Comparaison à des signatures génétiques connues

Nous avons mené une analyse d'enrichissement fonctionnelle au sein de nos 4 groupes de gènes d'intérêts : Les gènes différentiellement exprimés et les modules ■ brun, ■ vert et ■ turquoise (**Figure 11**). Pour obtenir un aperçu complet de l'enrichissement d'une famille de gènes dans un groupe de gènes d'intérêt, plusieurs paramètres sont souvent pris en compte. Le **GeneRatio** correspond au rapport entre le nombre de gènes du groupe d'intérêt qui appartiennent à une certaine famille de gènes et le nombre total de gènes dans ce groupe. Le **BgRatio** ou *Background Ratio* correspond au nombre de gènes dans la famille de gène en question par rapport au nombre total de gènes dans

l'analyse (également appelé "univers"). La **p-valeur** de l'enrichissement représente la probabilité que l'enrichissement observé soit dû au hasard. La **p-valeur ajustée (p.adjust)** est une version corrigée de la p-valeur d'enrichissement qui tient compte du nombre de tests statistiques effectués dans l'analyse. Tous ces paramètres nous permettent d'avoir une vision complète des phénomènes d'enrichissement.

L'analyse par sur-représentation des gènes différentiellement exprimés a révélé un enrichissement significatif de plusieurs familles de gènes exprimés dans des lymphocytes T naïfs, ainsi que de familles de gènes actifs dans d'autres types cellulaires en réponse au traitement par des cytokines (IFN γ et IFN α) (**Figure 11.A**). Ces mêmes familles sont également enrichies dans le module ■ brun (**Figure 11.B**).

L'analyse du module ■ vert a identifié seulement 8 familles de gènes enrichies (**Figure 11.C**). Parmi ces familles, 2 sont issues d'études portant sur les lymphocytes T et proviennent de la comparaison de lymphocytes T au repos avec des lymphocytes T activés.

Selon les résultats de l'analyse d'enrichissement fonctionnel du module ■ turquoise, la plupart des familles de gènes enrichies dans ce module n'ont pas de fonctionnalité commune évidente. (**Figure 11.D**). Une de ces familles est une famille de gènes de lymphocytes T épuisés retrouvés dans le modèle LCMV suite à une analyse par WGCNA (Doering et al., 2012).

(A)**Gènes différentiellement exprimés**

Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
GSE26495_NAIVE_VS_PD1HIGH_CD8_TCELL_UP	36/347	180/12865	1,61E-21	7,69E-18	7,10E-18
GSE26495_NAIVE_VS_PD1LOW_CD8_TCELL_UP	33/347	181/12865	1,66E-18	3,95E-15	3,65E-15
GSE45739_UNSTIM_VS_ACD3_ACD28_STIM_WT_CD4_TCELL_UP	31/347	169/12865	1,58E-17	2,51E-14	2,32E-14
GSE11057_NAIVE_VS_EFF_MEMORY_CD4_TCELL_UP	31/347	188/12865	3,70E-16	4,42E-13	4,08E-13
GSE5542_UNTREATED_VS_IFNA_TREATED_EPITHELIAL_CELLS_24H_DN	30/347	189/12865	3,30E-15	3,16E-12	2,91E-12
GSE11057_NAIVE_VS_MEMORY_CD4_TCELL_UP	28/347	176/12865	2,70E-14	2,15E-11	1,98E-11
GSE11057_NAIVE_VS_CENT_MEMORY_CD4_TCELL_UP	25/347	168/12865	3,20E-12	2,18E-09	2,01E-09
GSE5542_UNTREATED_VS_IFNG_TREATED_EPITHELIAL_CELLS_6H_UP	23/347	172/12865	2,21E-10	1,32E-07	1,22E-07
GSE22045_TREG_VS_TCONV_DN	22/347	160/12865	3,20E-10	1,70E-07	1,57E-07
GSE21670_IL6_VS_TGFB_AND_IL6_TREATED_CD4_TCELL_UP	22/347	162/12865	4,08E-10	1,95E-07	1,80E-07

(B)**Module brun**

Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
GSE26495_NAIVE_VS_PD1HIGH_CD8_TCELL_UP	72/332	180/12865	3,61E-68	1,71E-64	1,58E-64
GSE11057_NAIVE_VS_EFF_MEMORY_CD4_TCELL_UP	64/332	188/12865	4,15E-55	9,71E-52	8,97E-52
GSE26495_NAIVE_VS_PD1LOW_CD8_TCELL_UP	63/332	181/12865	6,16E-55	9,71E-52	8,97E-52
GSE11057_NAIVE_VS_MEMORY_CD4_TCELL_UP	60/332	176/12865	1,18E-51	1,40E-48	1,29E-48
GSE11057_NAIVE_VS_CENT_MEMORY_CD4_TCELL_UP	54/332	168/12865	7,30E-45	6,90E-42	6,38E-42
GSE5542_UNTREATED_VS_IFNA_TREATED_EPITHELIAL_CELLS_24H_DN	47/332	189/12865	2,53E-33	1,99E-30	1,84E-30
GSE5542_UNTREATED_VS_IFNG_TREATED_EPITHELIAL_CELLS_6H_UP	41/332	172/12865	2,44E-28	1,64E-25	1,52E-25
GSE5542_UNTREATED_VS_IFNG_TREATED_EPITHELIAL_CELLS_24H_DN	42/332	191/12865	1,70E-27	1,01E-24	9,30E-25
GSE5542_UNTREATED_VS_IFNA_AND_IFNG_TREATED_EPITHELIAL_CELLS_6H_UP	38/332	171/12865	4,04E-25	2,12E-22	1,96E-22
GSE5542_UNTREATED_VS_IFNA_TREATED_EPITHELIAL_CELLS_6H_DN	39/332	185/12865	7,49E-25	3,54E-22	3,27E-22

(C)**Module vert**

Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
GSE6674_ANTI_IGM_VS_ANTI_IGM_AND_CPG_STIM_BCELL_UP	16/286	187/12865	4,26E-06	1,12E-02	1,09E-02
GSE32986_UNSTIM_VS_GMCSF_STIM_DC_UP	16/286	193/12865	6,40E-06	1,12E-02	1,09E-02
GSE46606_UNSTIM_VS_CD40L_IL2_IL5_DAY1_STIMULATED_BCELL_UP	15/286	173/12865	7,25E-06	1,12E-02	1,09E-02
GSE9988_ANTI_TREM1_VS_CTRL_TREATED_MONOCYTES_DN	15/286	192/12865	2,53E-05	2,60E-02	2,53E-02
GSE21670_UNTREATED_VS_IL6_TREATED_STAT3_KO_CD4_TCELL_UP	14/286	171/12865	2,80E-05	2,60E-02	2,53E-02
GSE22886_NAIVE_CD4_TCELL_VS_12H_ACT_TH1_UP	14/286	184/12865	6,28E-05	4,68E-02	4,56E-02
GSE16450_CTRL_VS_IFNA_6H_STIM_IMMATURE_NEURON_CELL_LINE_DN	14/286	186/12865	7,06E-05	4,68E-02	4,56E-02
GSE5542_UNTREATED_VS_IFNA_TREATED_EPITHELIAL_CELLS_24H_DN	14/286	189/12865	8,39E-05	4,87E-02	4,74E-02

(D)**Module turquoise**

Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
GSE16450_CTRL_VS_IFNA_12H_STIM_IMMATURE_NEURON_CELL_LINE_UP	30/682	194/12865	1,13E-07	3,10E-04	2,91E-04
GSE21063_3H_VS_16H_ANTI_IGM_STIM_BCELL_DN	30/682	195/12865	1,27E-07	3,10E-04	2,91E-04
GSE17721_POLYIC_VS_GARDIQUIMOD_4H_BMDC_UP	26/682	171/12865	1,10E-06	1,79E-03	1,68E-03
GSE3982_MEMORY_CD4_TCELL_VS_TH1_UP	26/682	174/12865	1,55E-06	1,88E-03	1,77E-03
GSE22886_NAIVE_TCELL_VS_DC_UP	27/682	193/12865	3,57E-06	2,89E-03	2,72E-03
GSE23321_CD8_STEM_CELL_MEMORY_VS_NAIVE_CD8_TCELL_DN	27/682	193/12865	3,57E-06	2,89E-03	2,72E-03
GSE3982_MEMORY_CD4_TCELL_VS_TH2_UP	25/682	175/12865	5,60E-06	3,89E-03	3,65E-03
GSE41867_DAY6_EFFECTOR_VS_DAY30_MEMORY_CD8_TCELL_LCMV_ARMSTRONG_UP	26/682	190/12865	8,11E-06	4,63E-03	4,36E-03
GSE27241_WT_VS_RORGT_KO_TH17_POLARIZED_CD4_TCELL_TREATED_WITH_DIGOXIN_UP	24/682	168/12865	8,58E-06	4,63E-03	4,36E-03
GSE22886_NAIVE_CD4_TCELL_VS_12H_ACT_TH2_UP	25/682	184/12865	1,37E-05	6,65E-03	6,26E-03

Figure 11 : Analyse d'enrichissement fonctionnelle par sur-représentation. Tableaux des 10 familles de gènes de signalisation immunitaire (MSigDB) les plus enrichies au sein des gènes différentiellement exprimés (A), des modules brun (B), vert (C) et turquoise (D).

C. Corrélation entre l'expression des modules et l'expression de PD1 et TOX

Pour identifier les modules qui jouent un rôle dans l'épuisement des lymphocytes T CD8⁺, nous avons évalué la corrélation entre l'expression des *gènes propres* des différents modules et celle de *PD-1* et *TOX* (**Figure 12**). Celles-ci révèlent que les modules brun, vert et turquoise ne sont pas corrélés à l'expression de *PD-1* et *TOX*. De manière intéressante, le module lightcyan est positivement corrélé à l'expression de *PD-1*. Une revue de sa composition révèle qu'il contient *CTLA4* et *TIM3*, deux récepteurs inhibiteurs exprimés sur les T épuisés ainsi que IL-10, une cytokine anti-inflammatoire. Elle contient un lncRNA différentiellement exprimé mais celui-ci est complètement inconnu de la littérature.

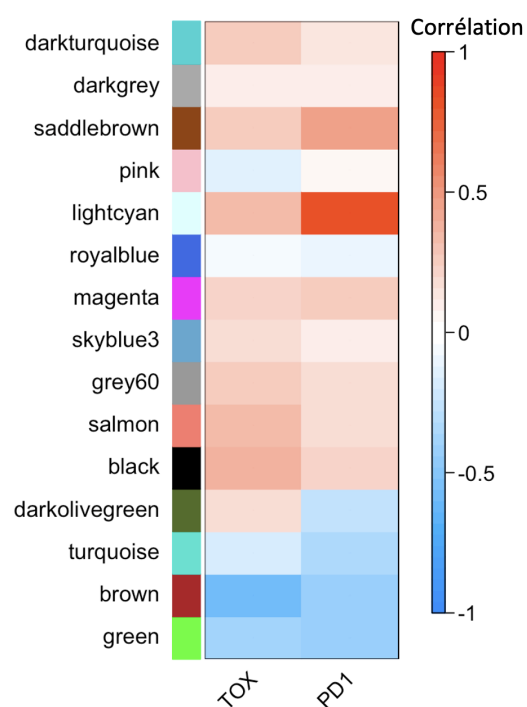


Figure 12 : Association entre les modules de gènes co-exprimés et l'expression de TOX et PD-1. Chaque ligne correspond au gène propre (ou tendance principale) d'un module. Chaque cellule représente la corrélation entre l'expression des gènes propres et l'expression de PD-1 ou TOX. Le code couleur correspond à l'intensité de la corrélation. Seuls les modules contenant au moins un des 94 lncRNAs différentiellement exprimés sont représentés.

4. Conclusions de l'analyse bio-informatique

Les modules que nous avons analysés plus en profondeur n'ont pas révélé des mécanismes connus typiques des lymphocytes T CD8⁺ épuisés. Au terme de notre analyse, nous avons choisi d'utiliser le lncRNA *LINC00402* pour commencer la mise au point de la caractérisation fonctionnelle. Le rôle de *LINC00402* a l'avantage d'avoir déjà été étudié dans des lymphocytes T humains. Selon cette étude, *LINC00402* stimule la signalisation induite par l'activation du TCR, ce qui entraîne une prolifération

soutenue et une augmentation de la production d'IL-2 (Peltier et al.,2021). Ce lncRNA présente de fortes chances de moduler les fonctions des lymphocytes T CD8⁺ et constitue donc le candidat parfait pour mettre au point notre méthode de caractérisation fonctionnelle de lncRNAs dans des lymphocytes T CD8⁺ primaires.

5. Caractérisation fonctionnelle des lncRNAs dans des lymphocytes T CD8⁺ humains.

5.1. Electroporation de siRNA

Pour inhiber l'expression du lncRNA LINC00402 dans des lymphocytes T CD8⁺ humains, nous avons d'abord essayé de nucléofecter des siARN. La nucléofection est une forme d'électroporation développée par AMAXA qui permet de transférer des molécules dans des cellules considérées comme "difficiles à transférer" en utilisant des paramètres électriques spécifiques pour provoquer des pores temporaires dans les membranes cellulaires. Le terme "nucléofection" vient du fait que cette méthode permet de transférer une molécule d'acide nucléique jusqu'au noyau de la cellule, ce qui permet de transférer des cellules quiescentes. À partir de maintenant et par souci de clarté, nous emploierons le terme "électroporation" au lieu de "nucléofection".

Nous avons électroporé des lymphocytes T CD8⁺ humains avec un siRNA ciblant CD2 ou un siRNA contrôle négatif (**Figure 13**). La protéine CD2 est une molécule d'adhésion cellulaire exprimée à la surface des lymphocytes T. CD2 est donc facilement observable par cytométrie en flux. Elle interagit avec LFA-3 et participe à la costimulation du TCR. Nous avons effectué de nombreux essais en variant plusieurs paramètres (concentrations de siRNA, congélation des cellules, activation du TCR, ...), et nous n'avons pu observer qu'une diminution moyenne de l'expression de CD2 de 10% (**Figure 13.B**).

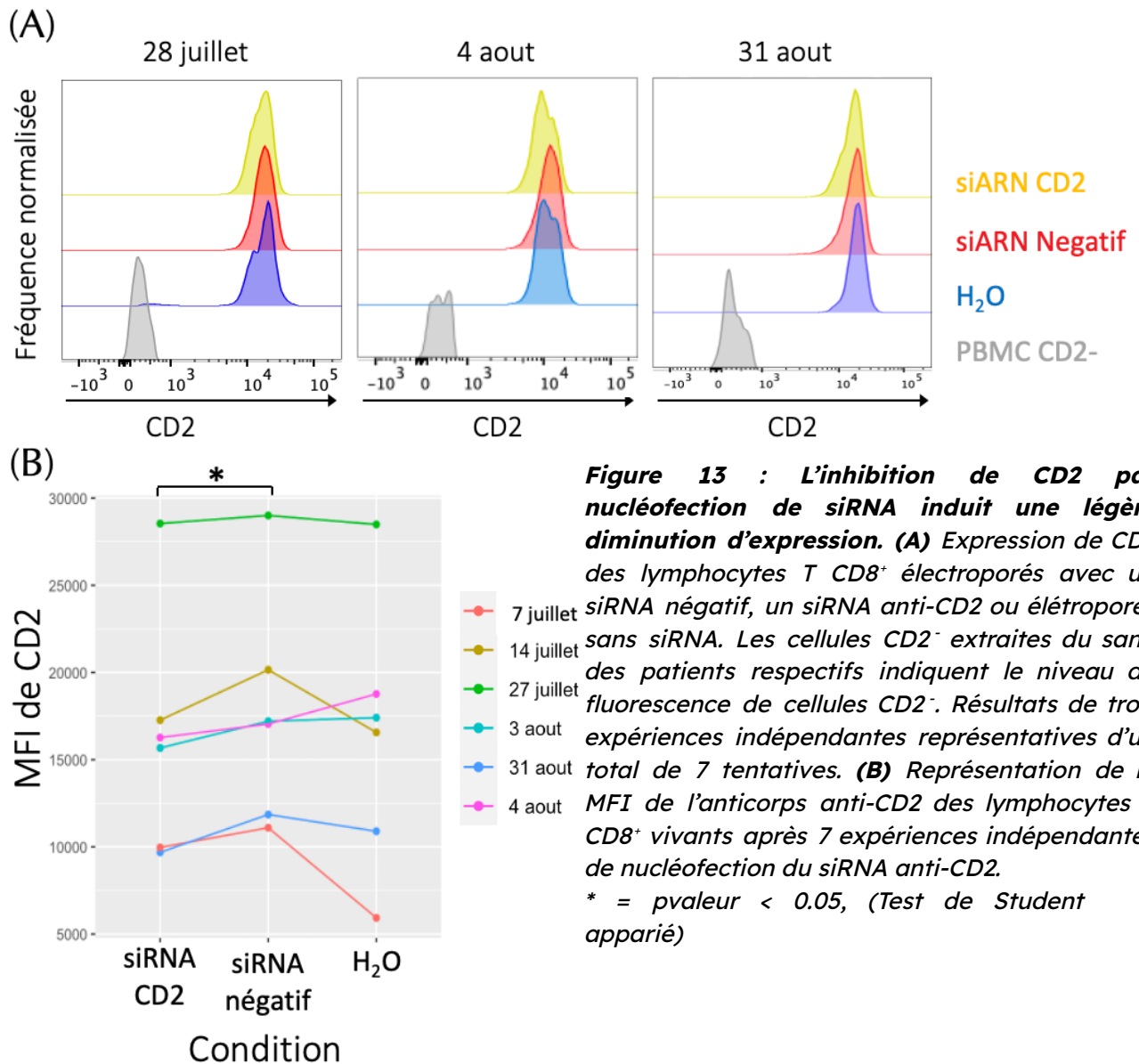


Figure 13 : L'inhibition de CD2 par nucléofection de siRNA induit une légère diminution d'expression. (A) Expression de CD2 des lymphocytes T CD8⁺ électroporés avec un siRNA négatif, un siRNA anti-CD2 ou électroporés sans siRNA. Les cellules CD2⁻ extraites du sang des patients respectifs indiquent le niveau de fluorescence de cellules CD2⁻. Résultats de trois expériences indépendantes représentatives d'un total de 7 tentatives. (B) Représentation de la MFI de l'anticorps anti-CD2 des lymphocytes T CD8⁺ vivants après 7 expériences indépendantes de nucléofection du siRNA anti-CD2.

* = p valeur < 0.05, (Test de Student apparié)

Pour tenter de comprendre l'origine de ces difficultés, nous avons élaboré l'expérience suivante (**Figure 14**). Des lymphocytes T CD8⁺ primaires au repos et activés ont été électroporés avec un siRNA anti-CD2 ou un siRNA anti-β2M dans trois types cellulaires : des lymphocytes T CD8⁺ au repos, des lymphocytes T CD8⁺ activés pendant 24h et des cellules HeLa. Pour chaque type cellulaire, une condition contrôle non-électroporée a également été mise en place afin de pouvoir comparer les résultats obtenus. Le siARN anti-β2M est un second contrôle positif pour nous assurer que l'absence de *knock-down* ne provient pas de l'inefficacité du siRNA anti-CD2. En comparant les données des lymphocytes T CD8⁺ primaires avec celles des cellules HeLa, connues pour être faciles à transfecter, nous pouvons nous assurer si c'est une spécificité de ces lymphocytes qui perturbe l'inhibition par siRNA. En comparant des lymphocytes T CD8⁺ au repos et

activés, nous pouvons également évaluer si l'activation a un impact sur l'efficacité du *knock-down*.

Enfin, chaque type cellulaire a également été électroporé avec un plasmide codant pour la GFP. Ce contrôle permet de vérifier l'efficacité d'électroporation sans impliquer la machinerie d'interférence ARN, qui a été décrite comme dysfonctionnelle dans les lymphocytes T CD8⁺ (Oberdoerffer et al., 2005 ; Mantei et al., 2008).

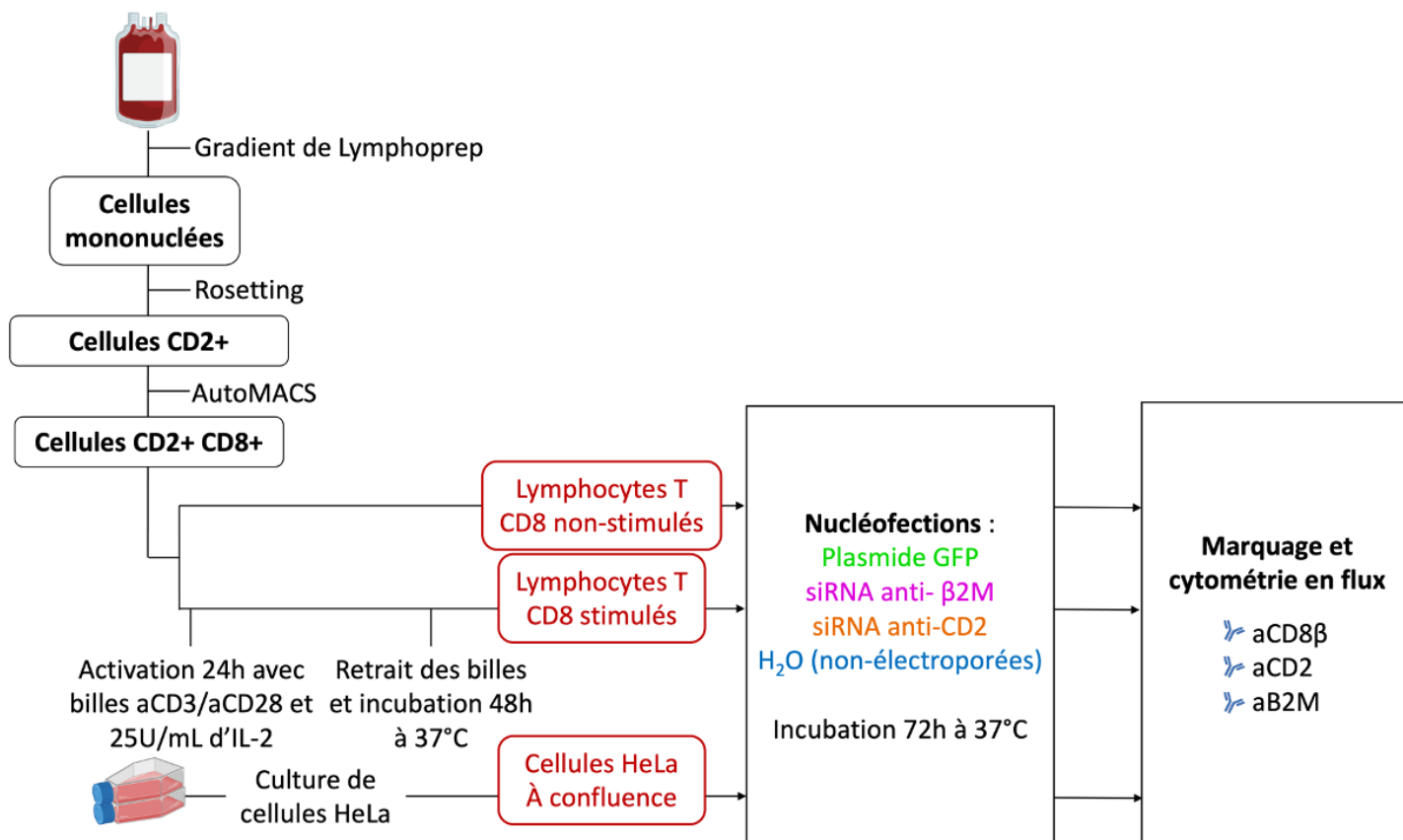


Figure 14 : Identifier l'origine des difficultés rencontrées lors de l'électroporation de siRNA. Représentation schématique de l'expérience.

Le plasmide codant pour la protéine GFP a été efficacement internalisé dans tous les types de cellules (**Figure 15.A**). Le siRNA anti-CD2 a provoqué une inhibition de 7 et 9% dans les lymphocytes T CD8⁺ (données non-présentées). De manière similaire, une inhibition de 10 à 12% de β2M a été observée dans les lymphocytes T CD8⁺ (**Figure 15.B**). Les cellules HeLa, quant à elles, présentent une inhibition de 66% de β2M. Ces données indiquent que l'origine de nos difficultés en matière d'électroporation de siRNA provient de la nature des lymphocytes T CD8⁺ et non du matériel utilisé.

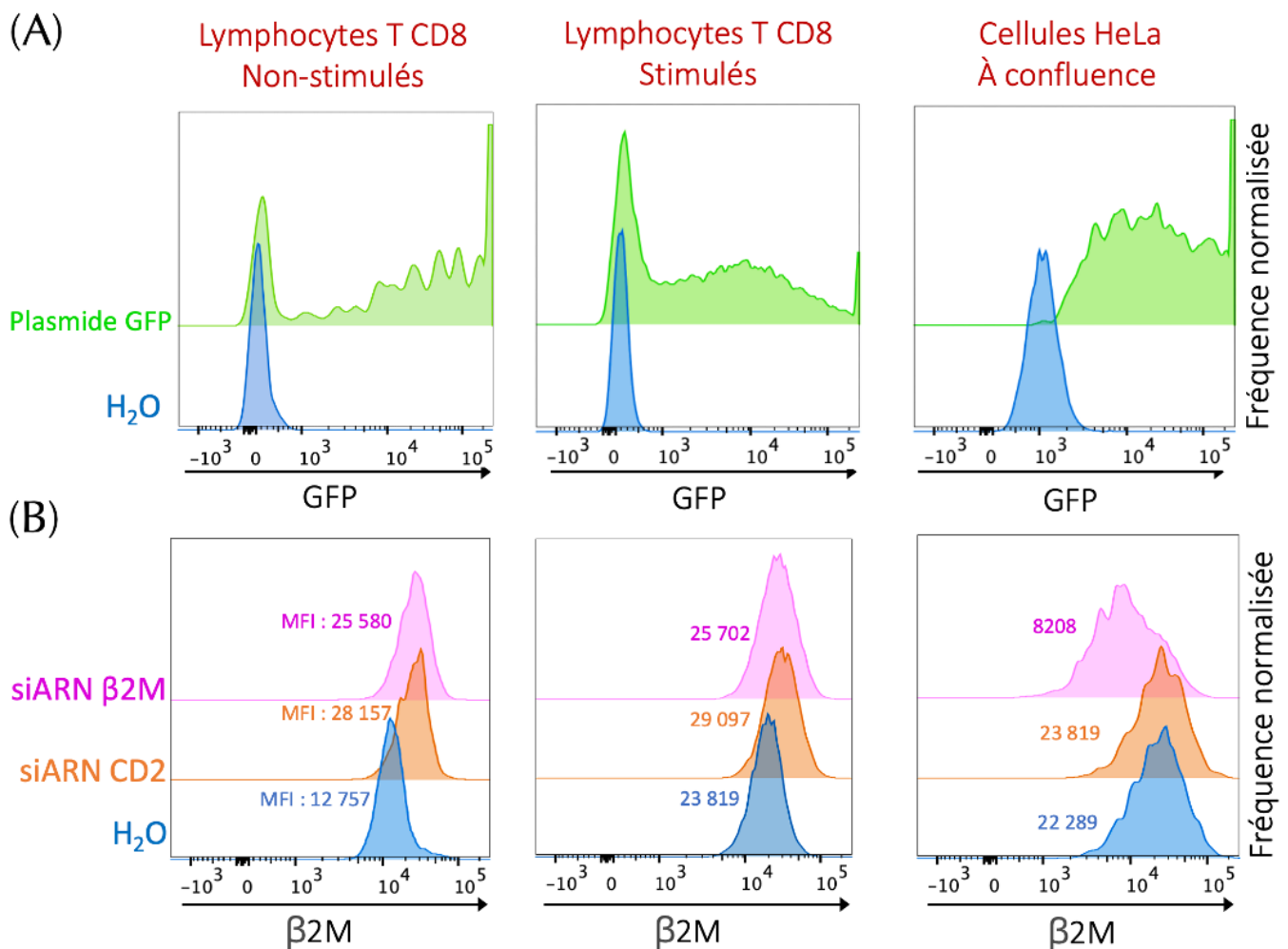


Figure 15 : L'inhibition de β 2M par électroporation de siRNA est efficace dans les cellules HeLa mais pas dans les lymphocytes T CD8⁺. (A) Contenus en GFP de différents types cellulaires 72 heures après avoir été électroporés avec un plasmide codant la GFP. (B) Expression de la β 2M par ces mêmes types cellulaires 72 heures après avoir été électroporés ou non avec un siRNA anti-CD2 ou un siRNA anti- β 2M. La fluorescence médiane (MFI) de chaque courbe est indiquée à la base de chaque courbe.

Suite à ces résultats et au vue de la littérature peu encourageante concernant l'électroporation de siRNA dans des lymphocytes T primaires (Oberdoerffer et al., 2005 ; Yin et al., 2006, Mantei et al., 2008), la décision a été prise de rechercher une autre approche pour inhiber l'expression des lncRNA.

5.2. Internalisation de GapmeR

Une brève explication au sujet des GapmeR est nécessaire pour comprendre les résultats obtenus. L'utilisation de GapmeR est une approche d'inhibition génique qui contourne les problèmes de mortalité causés par l'électroporation. Il s'agit de petits oligonucléotides antisens (*Antisens oligonucleotide, ASO*) constitués d'un fragment

d'ADN central de 5 à 10 nucléotides flanqués de deux fragments d'ARN modifiés chimiquement de 2 à 5 nucléotides, aussi appelés "*Locked Nucleic Acids*" ou LNA. **(Figure 16)**. Ils se prêtent particulièrement bien à notre étude pour plusieurs raisons. Premièrement, ils inhibent l'expression de leur ARN cible en s'y hybridant, ce qui crée un lien ADN-ARN capable d'induire son clivage par la RNase H. Le résultat de l'action des GapmeR est donc la dégradation de la molécule d'ARN cible (ce qui n'est pas systématiquement le cas pour l'interférence ARN) (Fazil et al., 2016). Par ailleurs, ils sont capables d'être internalisés dans les cellules par "gymnose", c'est-à-dire en l'absence de réactifs de transfection ou d'électroporation. En effet, il a été montré qu'ils peuvent être internalisés en profitant de la macropinocytose (Fazil et al., 2016). La macropinocytose est un processus qui permet l'absorption non sélective de petites molécules, telles que des bactéries, des virus ou des oligonucléotides. Elle est initiée par des replis de la membrane plasmique qui se replient sur eux-mêmes pour former des endosomes de tailles variables appelées macropinosomes. Cette internalisation a lieu spontanément dans les lymphocytes T. Par ce procédé, les risques de mortalité causés par l'électroporation ainsi que la résistance des lymphocytes T primaires aux agents de transfection sont évités (Yin et al., 2006, Mantei et al., 2008 ; Fazil et al., 2016).

Enfin, les modifications chimiques de leurs extrémités leur confèrent une forte affinité de liaison pour l'ARN cible, une meilleure internalisation et une résistance élevée aux nucléases qui améliore leur stabilité dans le sérum. Toutes ces caractéristiques permettent aux GapmeR solubilisés dans le milieu de culture d'induire une inhibition plus longue que leurs homologues siRNA.

L'efficacité des GapmeR a déjà été démontrée par le traitement de populations mixtes de lymphocytes T CD4⁺/CD8⁺ et de lymphocytes T uniquement CD4⁺ (Fazil et al., 2016 ; Abewe et al., 2020). Notre travail permet donc également d'évaluer la faisabilité de l'utilisation des GapmeRs dans les lymphocytes T CD8⁺ primaires humains.

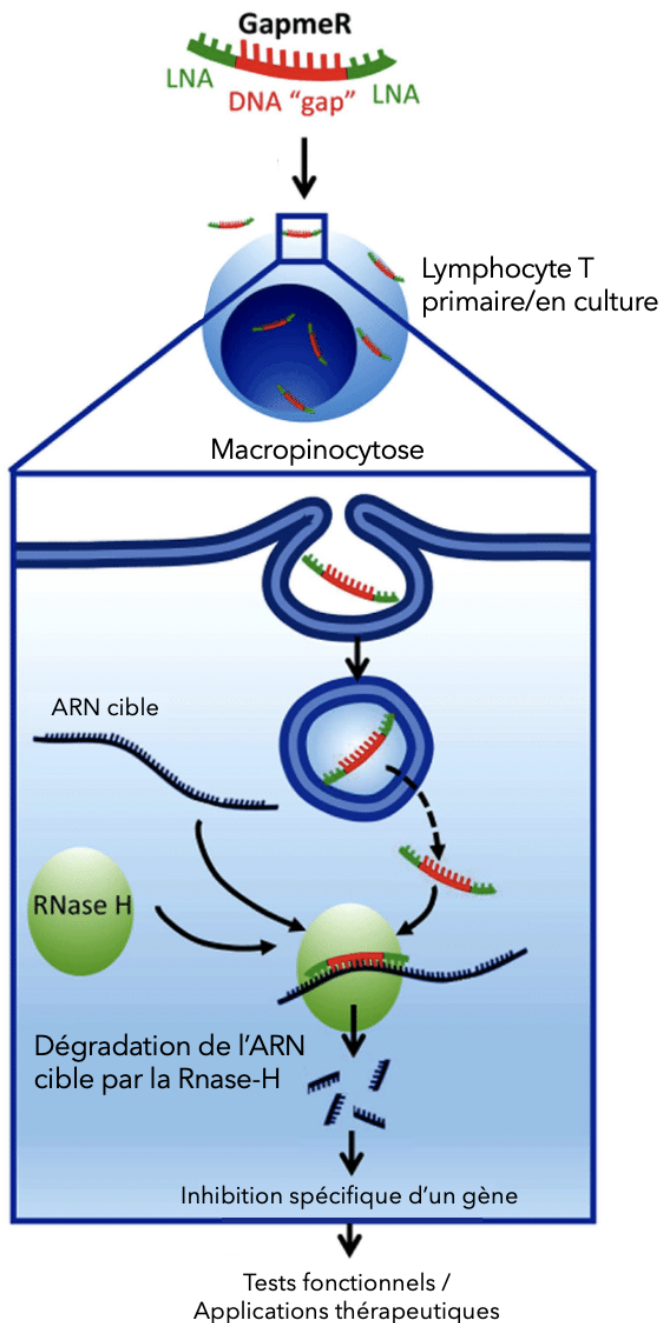


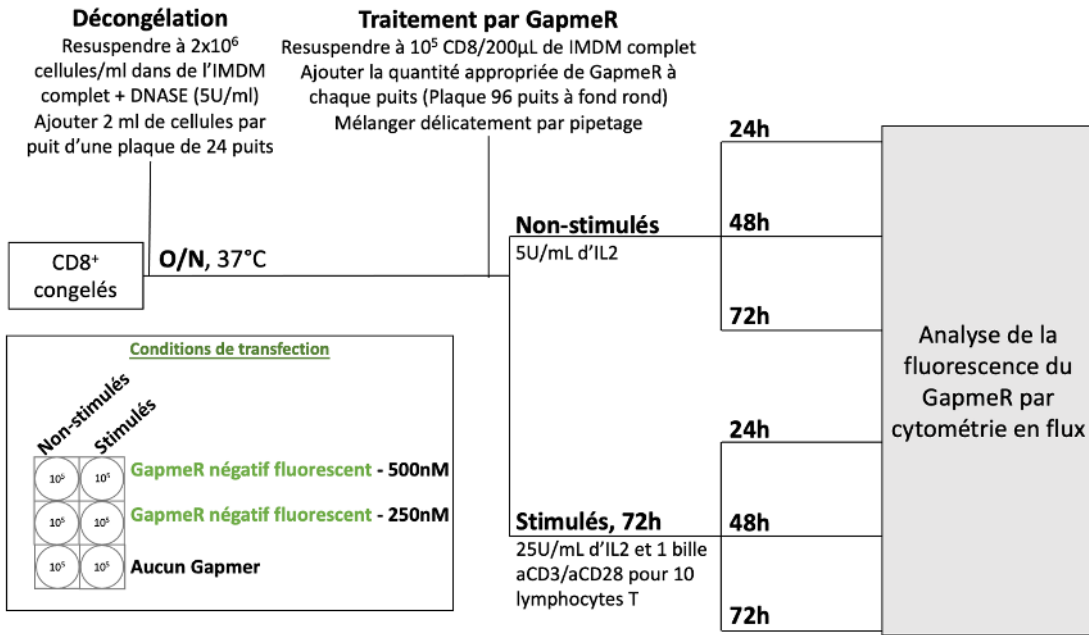
Figure 16 : Définition et mécanisme d'action des GapmeR. Les GapmeR sont de petits oligonucléotides constitués de deux fragments d'ARN modifiés chimiquement séparés par un écart (gap) fait d'ADN. Ils sont capables d'être internalisés par gymnose en s'incorporant dans les macropinosomes. Ils s'hybrident à leur ARN cible et crée un lien ADN-ARN capable d'induire le clivage par la RNase H. L'ARN cible est alors dégradé.

Adapté de Fazil et al., 2016

Afin d'évaluer la dynamique de l'internalisation des GapmeR dans les lymphocytes T CD8⁺ au cours du temps, nous avons cultivé des lymphocytes T

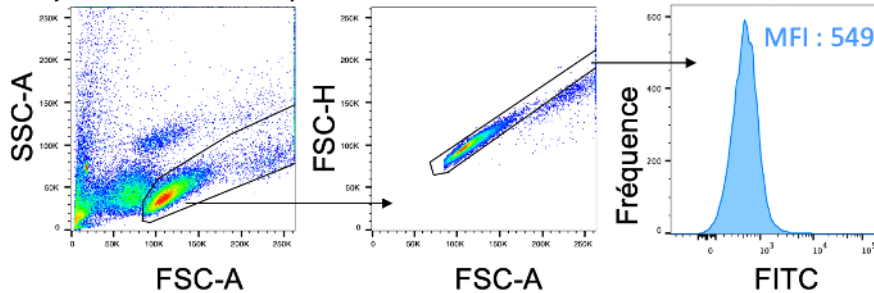
CD8⁺ primaires activés ou non avec plusieurs concentrations d'un GapmeR fluorescent pendant 24, 48 et 72 heures. Par ce biais, nous avons pu évaluer l'internalisation de ce GapmeR au cours du temps par cytométrie en flux (**Figure 17**). Les données confirment l'internalisation du GapmeR dès 24 heures. Elles indiquent également que les lymphocytes T CD8⁺ activés présentent une meilleure capacité d'internalisation que ceux qui sont au repos. Sur base de ces données et de la littérature (Fazil et al., 2016 ; Abewe et al., 2020), les conditions suivantes ont été choisies pour les expériences suivantes : une concentration de GapmeR de 500nM et une durée d'incubation de 72 heures.

(A)

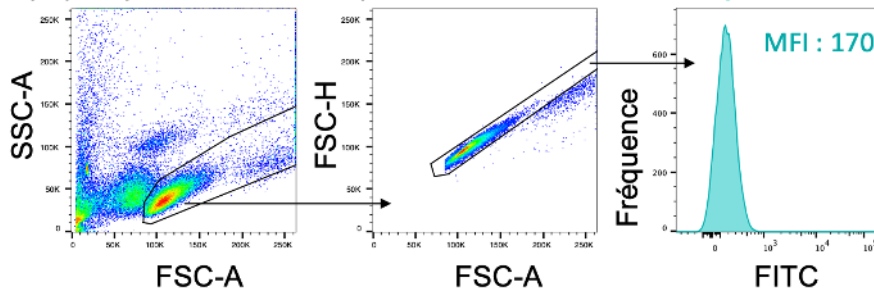


(B)

Lymphocytes T CD8 stimulés pendant 24h et traités avec 250nM de GapmeR Fluorescent



Lymphocytes T CD8 stimulés pendant 24h et traités sans GapmeR Fluorescent



(C)

Internalisation du GapmeR négatif fluorescent

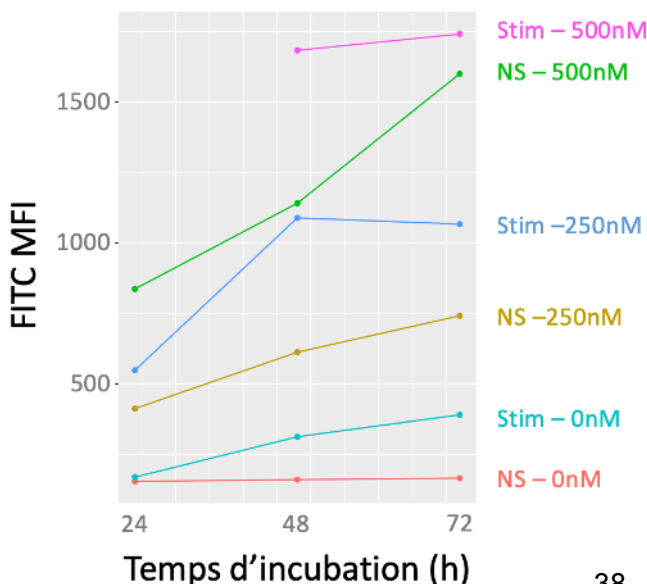


Figure 17 : Internalisation du GapmeR négatif fluorescent contrôle. Des lymphocytes T CD8+ humains ont été cultivés en présence de différentes concentrations de GapmeR fluorescent dans leur milieu et pendant 24, 48 et 72 heures. Parallèlement, ces lymphocytes ont été stimulés ou non par des billes couvertes d'anticorps anti-CD3 et anti-CD28. (A) Représentation schématique de l'expérience. (B) La stratégie de gating utilisée sélectionne les lymphocytes selon leur prodiffusion (FSC, forward-scatter) et leur diffusion latérale (SSC, side-scatter). Après avoir sélectionné les singulets, l'intensité médiane de fluorescence en FITC (530/30) est mesurée. (C) Evolution au cours du temps de la MFI en FITC des lymphocytes T CD8+ stimulés (Stim) ou non (NS) et traités avec différentes concentration d'un GapmeR fluorescent.

Nous avons testé l'efficacité de 3 GapmeR validés dans la littérature : deux GapmeR anti-CD11a et un GapmeR anti-*LINC00402* (Fazil et al., 2016, Peltier et al., 2021). CD11a est une protéine membranaire exprimée par les lymphocytes T CD8⁺ qui est ciblé par deux GapmeR validés dans la littérature (Ce qui n'était pas le cas de CD2). *LINC00402* est un lncRNA connu pour stimuler la signalisation du TCR et qui pourrait être impliqué dans l'épuisement. Des lymphocytes T CD8⁺ ont été traités avec ces GapmeR pendant 72 heures pour évaluer l'expression de leur ARN cible par RT-qPCR et l'expression protéique de CD11a par cytométrie en flux (**Figure 18**). Pour chaque condition, des lymphocytes T CD8⁺ stimulés ou non par des billes couvertes d'anticorps anti-CD3 et anti-CD28 (αCD3/αCD28) ont été étudiés.

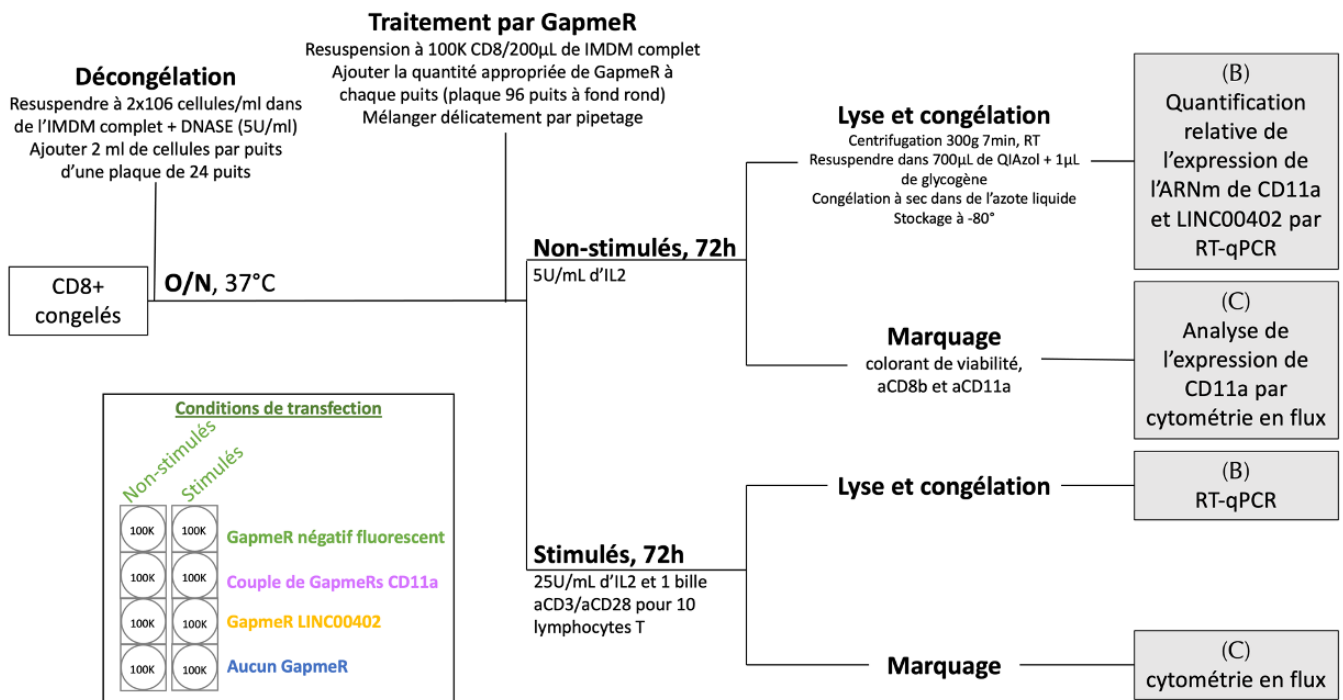


Figure 18 : Evaluation de l'efficacité des GapmeR. Représentation schématique de l'expérience

Les résultats de la RT-qPCR ont révélé un knock-down de 31 à 42% de l'expression des ARN ciblés par les GapmeR dans chacun des échantillons (**Figure 19.A**). A l'inverse, lorsque nous avons évalué l'expression de CD11a à la surface des lymphocytes T CD8⁺ traités par le couple de GapmeR anti-CD11a, aucune diminution de l'expression membranaire de CD11a n'a pu être observée (**Figure 19.B**).

De manière surprenante, on peut observer l'expression bimodale de CD11a dans les lymphocytes T CD8⁺ non-stimulés² (**Figure 19.B**).

(A) Quantification relative de l'expression de l'ARNm de CD11a et LINC00402 par RT-qPCR

Stimulation	GapmeR	ΔCq^*	$\Delta\Delta Cq^{**}$	Fold-change $2^{(-\Delta\Delta Cq)}$	% de Knock-down
Non	CD11	4,67	0,78	0,58	42
	Aucun	3,89			
Oui	CD11	4,51	0,67	0,63	37
	Aucun	3,84			
Non	LINC00402	5,98	0,53	0,69	31
	Aucun	5,45			
Oui	LINC00402	11,20	0,67	0,63	37
	Aucun	10,53			

* : ΔCq = cycle du gène d'intérêt - cycle du gènes de référence.

** : $\Delta\Delta Cq$ = ΔCq de l'échantillon traité - ΔCq de l'échantillon contrôle.

(B) Analyse de l'expression de CD11a par cytométrie en flux

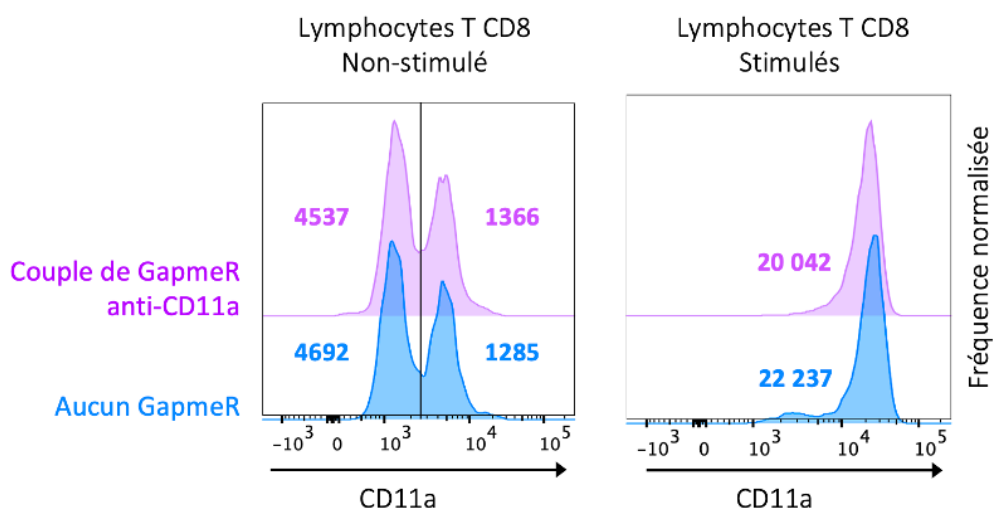


Figure 19 : Efficacités des GapmeR à inhiber l'expression de leur ARN cible et à inhiber l'expression membranaire de CD11a. Des lymphocytes T CD8⁺ décongelés ont été stimulés ou non et traités avec différents GapmeR pendant 72 heures. (A) Quantification relative par RT-qPCR du niveau d'expression de CD11a et de LINC00402, 72 heures après l'ajout des GapmeR. Le gène de référence utilisé est EF1. (B) Evaluation de l'effet du couple de GapmeR anti-CD11a sur l'expression membranaire de CD11a par cytométrie en flux. La fluorescence médiane (MFI) est indiquée pour chaque pic d'expression.

² Cette observation est retrouvée dans la littérature et plusieurs explications sont avancées (Desroches et al., 1989, Barrena et al., 1992, Capece, 2016). Ces deux populations expriment certains récepteurs de chémokines à des niveaux différents et diffèrent dans leur capacité à être retenu dans les ganglions lymphatiques (Capece, 2016). La littérature s'accorde pour dire que ces deux populations sont liées à des différences fonctionnelles (Desroches et al., 1989, Barrena et al., 1992, Capece, 2016).

5.3. Impact de l'inhibition de LINC00402 sur les fonctions de lymphocytes T CD8

Nous avons évalué l'impact de l'inhibition de LINC00402 sur les fonctions des lymphocytes T CD8⁺. Pour cela, nous avons incubé des lymphocytes T CD8⁺ pendant 72 heures avec un GapmeR ciblant *LINC00402* avant de les soumettre à différents tests de fonction : un test de prolifération et des tests de sécrétion de cytokine par ELISA (**Figure 20.A**).

Pour réaliser les tests de sécrétion de cytokines, nous avons prélevé le surnageant de lymphocytes T CD8⁺, activés ou non, pendant 20 heures par des billes α CD3/ α CD28. Nous avons utilisé ce surnageant pour déterminer la quantité d'IFN γ et de TNF α sécrétée par les lymphocytes T CD8⁺. Les résultats obtenus avec les deux tests ELISA sont similaires : il semble que le GapmeR ciblant LINC00402 diminue la sécrétion d'IFN γ et de TNF α par les lymphocytes T CD8⁺ après 72 heures (**Figure 20.B**).

Pour réaliser le test de prolifération, nous avons activé des lymphocytes T CD8⁺ pendant 5 jours après les avoir marqués au Cell Trace Violet. Ce colorant fluorescent est incorporé dans les cellules et lie de manière permanente les protéines cytosoliques. Ce colorant est dilué par deux à chaque division. Le niveau de fluorescence des cellules marquées permet alors de déterminer le nombre de divisions qu'elles ont subies (Lemieszek et al ., 2022).

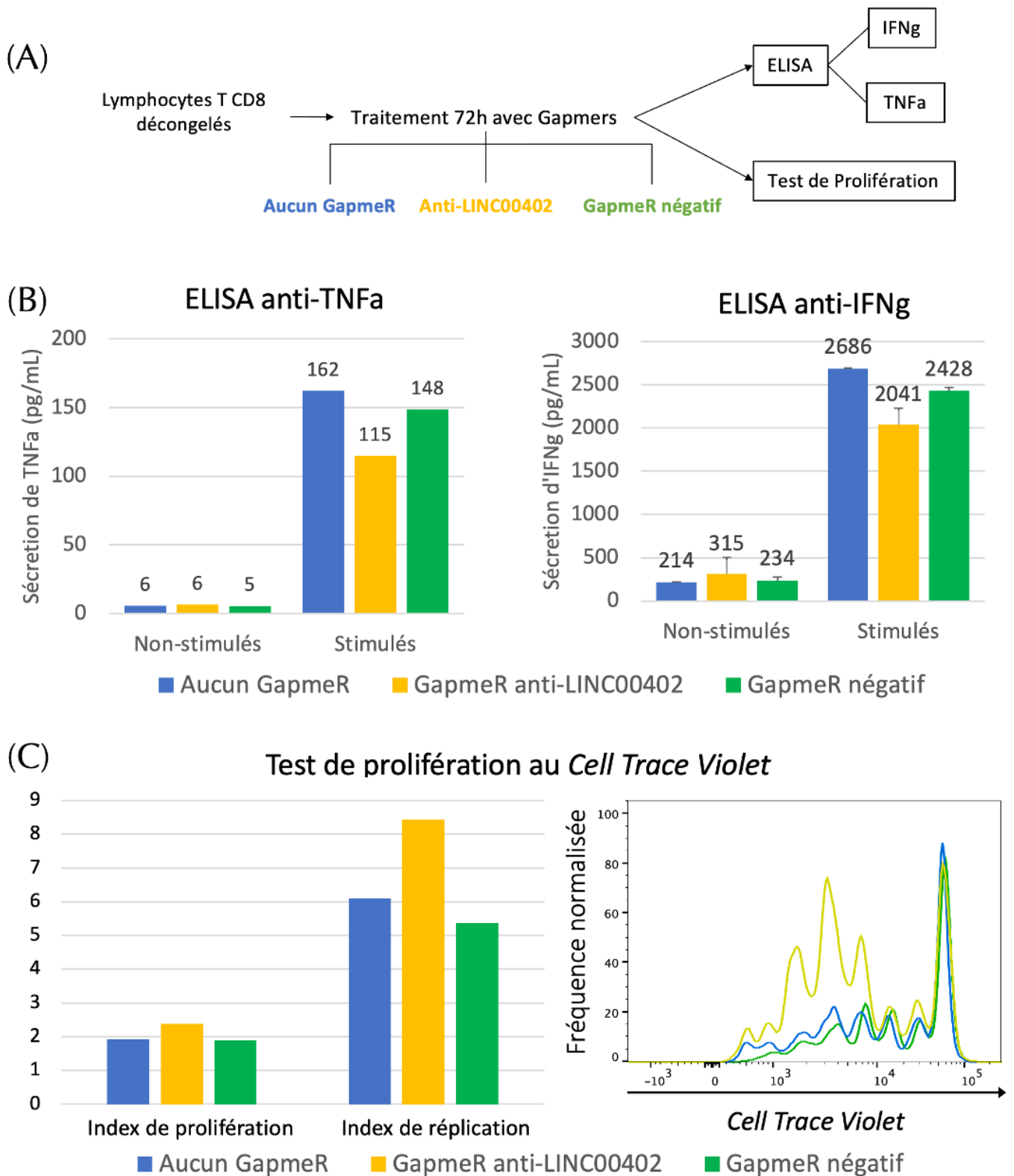


Figure 20 : Le GapmeR anti-LINC00402 provoque une baisse de la sécrétion d'IFN γ et de TNF α ainsi qu'une augmentation de la prolifération des lymphocytes T CD8 $^+$. (A) Plan expérimental (B) Quantité de cytokines sécrétées dans le surnageant des lymphocytes T CD8 $^+$ déterminé par ELISA. (C) Résultats de l'analyse par cytométrie en flux du contenu en Cell Trace Violet des lymphocytes T CD8 $^+$ soumis au test de prolifération. (Gauche) Index de prolifération et index de répliation des différents échantillons. (Droite) Fluorescence des différents échantillons pour le détecteur du CTV (450/40).

L'analyse de la fluorescence du colorant *Cell Trace violet* est résumée par deux index. L'index de prolifération est calculé en divisant le nombre total de divisions par le nombre de cellules qui ont subi une division. L'index de réplication est calculé en divisant le nombre total de cellules qui ont subi une division par le nombre de cellules qui ont subi une division. Ces deux index sont utilisés pour représenter la prolifération et révèlent que l'inhibition de LINC00402 stimule la prolifération des lymphocytes T CD8⁺ (**Figure 20.C**).

En résumé, il semble que l'utilisation de GapmeR soit une approche capable d'inhiber l'expression de lncRNAs dans les lymphocytes T CD8⁺ et d'induire des changements fonctionnels sans provoquer de mortalité ni d'activation du TCR.

Discussion

1. Préambule

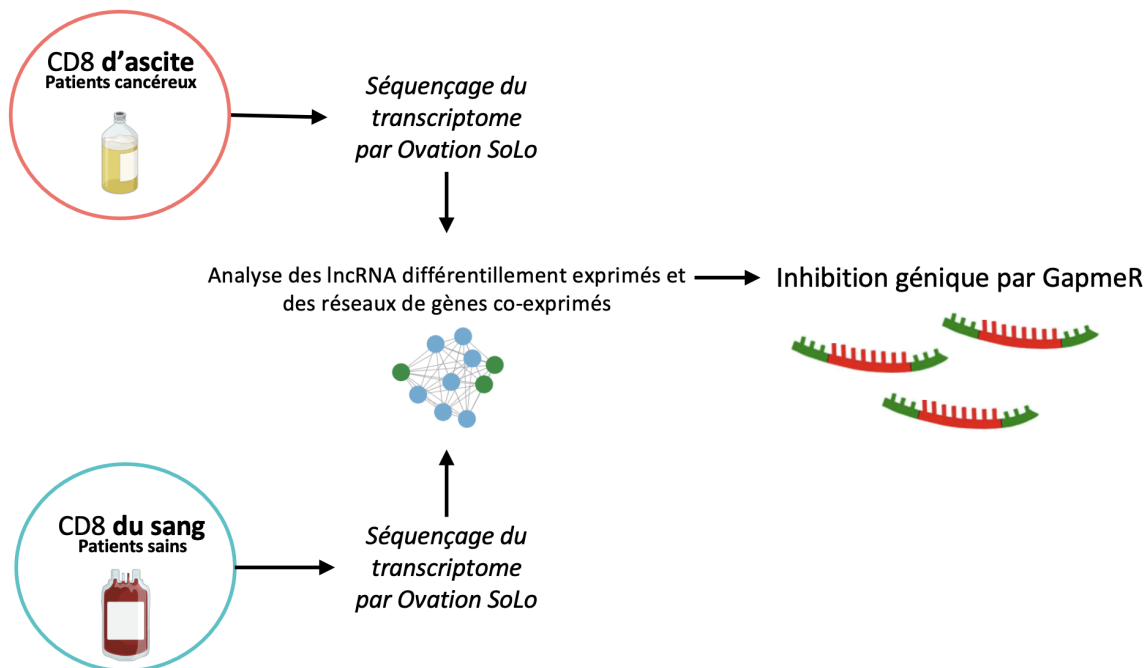


Figure 21 : Objectifs et résultats de ce mémoire. Les lymphocytes T CD8⁺ isolés à partir d'ascite carcinomateuses et du sang de donneurs non-cancéreux ont été séquencées avec le kit Ovation SoLo. Ces données ont été utilisées pour élaborer une analyse bioinformatique visant à identifier les lncRNAs différemment exprimés et à analyser les réseaux de gènes co-exprimés. À partir d'un lncRNA identifié dans cette analyse, l'inhibition de l'expression de lncRNAs par GapmeR a été mise au point dans des lymphocytes T CD8⁺ humains. Réalisé avec BioRender.com

Nous avons utilisé des données transcriptomiques de lymphocytes T CD8⁺ d'ascites tumorales et du sang de donneurs non-cancéreux pour mettre au point une approche capable d'identifier des lncRNAs impliqués dans l'épuisement des lymphocytes T CD8⁺ (**Figure 21**). Notre analyse bioinformatique nous a permis d'identifier 3 lncRNAs qui pourraient être impliqués dans l'épuisement des TILs selon la littérature : LINC00402, SATB1-AS1 et LINC00649. Nous avons ensuite voulu développer une approche pour tester leur fonction dans les lymphocytes T CD8⁺. Pour cela, nous avons d'abord tenté de modifier le niveau d'un gène par l'électroporation de siRNA dans ces cellules, mais cela s'est révélé très difficile. Nous nous sommes alors tournés vers l'utilisation de GapmeR et avons ainsi réussi à inhiber l'expression de LINC00402 à hauteur de 30%. Enfin, nous avons vérifié l'effet de cette inhibition sur la fonction des lymphocytes T CD8⁺ humains et avons constaté une réduction de la production d'IFN γ et de TNF α ainsi qu'une augmentation de la prolifération.

2. Est-ce que nous avons le matériel adéquat pour répondre à nos objectifs ?

Comme indiqué dans l'introduction (page 13), notre jeu de données de lymphocytes T CD8⁺ d'ascites tumorales et sanguins n'a pas été conçu dans le but de répondre à nos objectifs de recherche. L'analyse d'enrichissement fonctionnel des gènes différentiellement exprimés confirme que les signatures transcriptomiques des lymphocytes T naïfs constituent une différence majeure dans nos données (**Figure 11.A, page 30**). Malgré cela, ce jeu de données nous a permis de découvrir une série de lncRNAs pour lesquels la littérature suggère un rôle dans les fonctions des lymphocytes T CD8⁺. De plus, notre analyse met en évidence l'association de ces lncRNAs avec divers processus tels que l'activation du TCR ou la quiescence.

3. Qu'est-ce que l'analyse des réseaux de co-expression a apporté de plus que l'analyse de la littérature ?

L'analyse classique des gènes différentiellement exprimés nous a révélé 94 lncRNAs dont 11 ont été identifiés comme intéressants suite à l'analyse de la littérature. L'analyse d'enrichissement fonctionnel des gènes différentiellement exprimés a révélé l'enrichissement de gènes spécifiques des lymphocytes T naïfs, de gènes liés à l'activation des lymphocytes T suite à une stimulation *in vitro*, et de gènes exprimés par

les cellules endothéliales en réponse à l'IFN γ et à l'IFN α . Cela suggère que les différences majeures entre les lymphocytes T CD8 $^+$ d'ascites tumorales et du sang de donneurs non-cancéreux sont leur état d'activation et leur proportion de lymphocytes T CD8 $^+$ naïfs.

L'analyse des réseaux de co-expression par WGCNA permet d'identifier les réseaux de gènes co-exprimés. Malgré le caractère strictement corrélatif de ces modules, leur composition indique souvent la participation de gènes à des processus biologiques similaires. (Albert, 2005, van Dam et al., 2018). Leur analyse peut donc nous renseigner sur la fonction de ces gènes, qu'ils soient différentiellement exprimés, codants ou non-codants.

L'analyse du réseau vert a notamment permis, en plus d'identifier LINC00402 comme potentiellement exprimés par les lymphocytes T naïfs, de découvrir son association avec une série de gènes impliqués dans le maintien de la quiescence. De plus, les gènes les plus coexprimés à LINC00402 sont TRBV6-1, un composant du TCR, et ITGAM, une intégrine qui participe à l'adhésion cellulaire et l'adhérence aux endothéliums. Il serait donc intéressant d'étudier le niveau d'expression de ces protéines ainsi que les capacités d'adhésion cellulaires suite à l'inhibition de LINC00402 dans des lymphocytes T CD8 $^+$ activés ou non. C'est particulièrement pertinent d'identifier des partenaires potentiels afin d'aller plus loin dans la caractérisation fonctionnelle des lncRNA.

4. Peut-on déduire des fonctions biologiques des 3 modules de gènes co-exprimés étudiés dans ce mémoire ?

Les trois modules partagent une tendance d'expression similaire : ils sont sous-exprimés dans l'ascite, sur-exprimés dans le sang et sont négativement corrélés à l'expression de PD1 et TOX (**Figures 9 et 12**). Ces résultats coïncident avec une conclusion commune de l'analyse des fonctions de ces 3 modules : ils semblent impliqués dans les fonctions des lymphocytes T naïfs.

Les fonctions connues des gènes hubs du module ■ brun et les familles de gènes qui y sont enrichies suggèrent que les gènes de ce module soient des gènes qui participent

au maintien de la quiescence contrôlé par *LEF1* et à la coordination du homing aux différents organes lymphoïdes par *CD62L* (**Figure 11.B, page 30**). Plusieurs des gènes hubs semblent participer directement à la voie Wnt et il se pourrait donc que ce soit cette voie de signalisation qui unit les gènes du module brun.

Le module ■vert, quant à lui, réunit des gènes liés à l'activation et à la sortie des lymphocytes T naïfs de leur quiescence. En effet : *RIPOR2* participe au maintien de la quiescence des lymphocytes T naïfs et est fortement co-exprimé aux autres gènes hubs (**Figure 11.C, page 30**). *TRAF3IP3* et *CUL4B* régulent la signalisation de mTORC1, qui a été décrit comme un acteur majeur de la sortie de quiescence des lymphocytes T naïfs (Chapman et Chi., 2018). De plus, les gènes *SDCCAG8*, *DNASE1*, *MAP2K6* et *ABI2* présentent tous une fonction qui soutient la prolifération cellulaire.

Le module ■turquoise a une tendance principale strictement sur-exprimée dans le sang et sous-exprimée dans l'ascite. La littérature de ses gènes hubs et les familles de gènes les plus enrichies en son sein ne révèlent pas de fonction commune évidente. Notre analyse ne nous permet pas d'identifier une fonction biologique commune des gènes du module turquoise. Néanmoins, compte tenu de la tendance principale du module et du rôle de son meilleur gène hub, *RAC2*, dans l'activation du TCR, on peut suggérer que ce module de gènes soit lié à l'activation des lymphocytes T CD8⁺ dans l'ascite (**Figures 9-12**).

5. Quelles sont les limitations de l'analyse des réseaux de co-expression par WGCNA?

L'algorithme de "*Dynamic branch cutting*" permet de construire des réseaux de co-expression de gènes en découpant de manière dynamique les branches du dendrogramme de co-expression. Les paramètres de cet algorithme sont utilisés pour déterminer comment les branches des réseaux sont découpées et, par conséquent, comment les réseaux de co-expression de gènes sont construits. Ces paramètres sont choisis de manière arbitraire, en fonction des objectifs de l'analyse. L'utilisateur peut ainsi varier la taille, le nombre ou l'allure des modules en variant la sensibilité de l'algorithme pour, par exemple, fusionner deux modules qui partagent la même annotation fonctionnelle ou diminuer le nombre de clusters fortuits dans une analyse

qui ne comprend pas un nombre recommandé d'échantillons. Cet aspect arbitraire rend important de varier les paramètres pour s'assurer de la robustesse des réseaux observés, c'est-à-dire leur capacité à ne pas varier significativement en fonction des paramètres choisis. Il est donc recommandé de tester les réseaux en variant les paramètres pour s'assurer de la fiabilité des résultats. C'est ce qui a été fait dans cette analyse pour choisir les paramètres qui ont permis de construire nos réseaux de co-expressions (**Annexe 1, page 69**).

Il y a deux types majeurs de réseaux de co-expression : les réseaux signés et les réseaux non-signés. Dans les réseaux signés, les corrélations entre l'expression des gènes se trouvent entre -1 et +1. Ce ne sont que les gènes les plus positivement corrélés (c'est-à-dire que les gènes sont exprimés de manière similaire) qui y sont considérés comme co-exprimés. Ces types de réseaux sont souvent mieux séparés et plus significatifs. Dans les réseaux non-signés, les corrélations entre l'expression des gènes se trouvent entre 0 et 1, peu importe le signe du coefficient de corrélation. Les gènes co-exprimés dans de tels réseaux partagent soit une corrélation positive, soit une corrélation négative (c'est-à-dire que les gènes sont exprimés de manière inversement proportionnelle). Cette approche permet de mieux inclure les gènes qui antagonisent l'expression d'autres gènes, comme les microARN ou les lncRNA. C'est pour cette raison que nous avons construit des réseaux non-signés dans le cadre de ce travail de mémoire, malgré leur caractère sous-optimal.

Une limitation majeure de l'analyse de co-expression par WGCNA est le nombre minimal d'échantillons. Il n'y a pas de limite fixe mais un nombre d'au moins 15 échantillons est conseillé. Si cela est possible, il est recommandé d'avoir au moins 20 échantillons. Notre jeu de données (16 échantillons) est donc à peine suffisant pour donner des résultats convaincants. Plus d'échantillons conduisent généralement à des résultats plus robustes.

6. Pourquoi les lymphocytes T CD8⁺ sont-ils si difficiles à transfecter par siRNA ?

Il semble raisonnable de dire, au terme de nos expériences d'électroporation, que l'utilisation de siRNA pour inhiber un gène dans les lymphocytes T CD8⁺ primaires humains ne semble pas une approche efficace. Ces difficultés sont explicables par plusieurs paramètres. Les lymphocytes T sont des cellules primaires qui sont connues pour "être difficile à transfecter" (Yin et al., 2006). De plus, ils ont été décrits comme possédant une machinerie d'interférence ARN moins efficace que d'autres cellules du sang, notamment à cause d'une dégradation plus importante des siRNA intra-cellulaires. (Oberdoerffer et al., 2005 ; Mantei et al., 2008). Il semble même que l'activation des lymphocytes T diminue drastiquement la demi-vie des siRNA conventionnels intracellulaires (Mantei et al., 2008). Ce sont pour toutes ces raisons que nous avons décidé d'utiliser une autre approche d'inhibition génique.

7. Comment interpréter l'absence d'inhibition de CD11a à la surface des lymphocytes T CD8⁺ traités avec les GapmeR anti-CD11a ?

Bien que le traitement par les GapmeR ait entraîné une diminution de 40% de la quantité d'ARNm de CD11a dans les lymphocytes T CD8⁺ humains, cela n'a pas eu un tel impact sur l'expression de la protéine CD11a (**Figure 19, page 40**). Il n'y a pas d'explication claire et évidente à ce constat. Cela pourrait s'expliquer par le fait que, selon la littérature, CD11a est stockée dans des endosomes pour être mobilisée en cas d'activation (Capece et al., 2017). Il est donc possible que ces endosomes aient fourni une source de CD11a membranaire déjà traduit, malgré la baisse de traduction induite par l'action des GapmeR. Dans les lymphocytes T CD8⁺ stimulés, l'expression massive de CD11a membranaire en réponse à l'activation pourrait contrecarrer l'inhibition causée par les GapmeR. On peut néanmoins constater une différence d'environ 10% de la MFI de CD11a dans cette condition. Il faudrait recommencer cette expérience pour s'assurer de la significativité d'une telle différence.

8. Quel est l'impact de l'inhibition de LINC00402 dans des lymphocytes T CD8⁺ humains ?

Dans l'article de l'équipe de Daniel Peltier, LINC00402 a été identifié comme un lncRNA impliqué dans l'allo-immunité des patients souffrant de GVHD (Graft versus Host Disease, maladie du greffon contre l'hôte) (Peltier et al., 2021). Dans cet article, des lymphocytes T humains sont traités 72 heures avec le GapmeR dirigé contre LINC00402 utilisé dans ce mémoire. Suite à ce traitement, le niveau de LINC00402 baisse de façon comparable à ce que nous avons observé (Entre 30 et 35%). En surexprimant LINC00402 dans des cellules Jurkat, ils ont pu montrer que LINC00402 stimule la sécrétion d'IL-2 induite par la signalisation ERK/FOS qui découle de l'activation du TCR. De plus, il induit une baisse de la prolifération des lymphocytes T. De manière étonnante, cet effet sur la prolifération n'est observé qu'en réponse à un stimulus allogénique (des monocytes isolés à partir d'autres patients) et non en réponse à l'activation par des billes α CD3/ α CD28. C'est, entre autres, pour ces raisons que LINC00402 est décrit comme un régulateur de l'allo-immunité. Néanmoins, cet article montre que la surexpression de LINC00402 stimule l'activité des voies de signalisation ERK/FOS peu importe le type de stimulus qui active le TCR. Il suggère donc que LINC00402 pourrait être impliqué dans les fonctions lymphocytaires au-delà des réponses allo-immunes.

Nos résultats sont en accord avec cette suggestion car l'activation par les billes α CD3/ α CD28 de lymphocytes T CD8⁺ traités de manière similaire a conduit à une diminution de la sécrétion d'IFN γ et de TNF α de 15% et 10% respectivement et à une augmentation de la prolifération (**Figure 20, page 42**). Ces résultats suggèrent une action contradictoire de LINC00402 sur la prolifération. En effet, les résultats de Peltier et al suggèrent que LINC00402 stimule la prolifération lymphocytaire et nos résultats suggèrent l'inverse (Peltier et al., 2021). L'action de LINC00402 sur la prolifération dépendrait donc du type de stimulus qui active les lymphocytes T. Il convient de souligner que nos résultats sont issus d'une seule expérience et qu'il sera nécessaire de les répéter pour établir l'impact réel de l'inhibition de LINC00402 sur les fonctions des lymphocytes T CD8⁺.

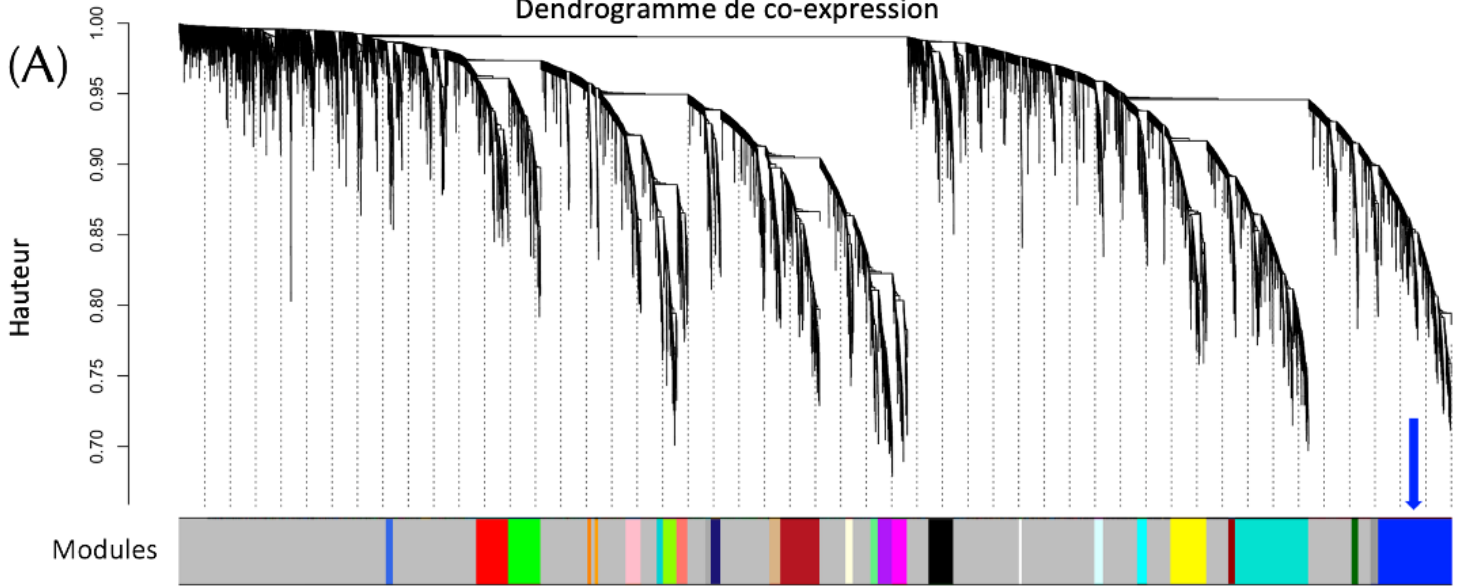
Il est important de souligner que selon l'article de Daniel Peltier, la quantité d'expression de LINC00402 dans les lymphocytes T humains activés par des billes α CD3/ α CD28 diminue rapidement après l'activation. Après 48 heures d'activation, moins de 1/100 de la quantité de LINC00402 persiste dans les lymphocytes T. Nos résultats de RT-qPCR obtenus montrent que le niveau de LINC00402 diminue de 97% après 72 heures de culture de lymphocytes T CD8⁺ stimulés ou non, ce qui est en accord avec la littérature (données non-présentées). Les tests fonctionnels ont été réalisés en stimulant des lymphocytes T CD8⁺ pendant 20 et 120 heures, on peut donc supposer que l'impact de GapmeR anti-LINC00402 sur les fonctions des lymphocytes T CD8⁺ s'est produit au début du processus d'activation. Toutes ces données suggèrent que LINC00402 joue un rôle dans la régulation précoce de l'activation des lymphocytes T CD8⁺.

9. Est-ce que ma méthode bioinformatique est applicable à d'autres données ?

L'analyse des réseaux de gènes co-exprimés a été appliquée à un autre jeu de données du laboratoire : les transcriptomes d'échantillons de lymphocytes T CD8⁺ générés par Thibault Hirsch dans l'étude du facteur de transcription IRF4 (**Figure 21.A, page 51**).

IRF4 a été identifié comme étant impliqué dans la voie NF- κ B induite par l'activation du TCR (Hirsch, 2022). L'analyse de sur-représentation des familles de gènes enrichies dans le module qui contient IRF4, le module ■bleu, révèle un enrichissement significatif de la famille des gènes induits par la signalisation NF- κ B (**Figure 21.B, page 51**). La tendance principale du module ■bleu, quant à elle, soutient qu'IRF4, et les gènes qui y sont co-exprimés, sont surexprimés en réponse à l'activation du TCR (**Figure 21.C, page 51**). L'analyse des réseaux de gènes co-exprimés confirme les résultats obtenus précédemment et l'étude des gènes co-exprimés avec IRF4 pourrait être utilisée pour identifier des gènes régulés par IRF4 qui sont encore inconnus.

Dendrogramme de co-expression



(B)

Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue	Count
HALLMARK_TNFA_SIGNALING_VIA_NFKB	48/287	172/3251	4,38E-14	2,19E-12	1,89E-12	48
HALLMARK_MYC_TARGETS_V2	15/287	58/3251	9,65E-05	2,41E-03	2,08E-03	15
HALLMARK_INFLAMMATORY_RESPONSE	25/287	136/3251	2,49E-04	4,15E-03	3,58E-03	25
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	21/287	112/3251	5,97E-04	7,47E-03	6,44E-03	21

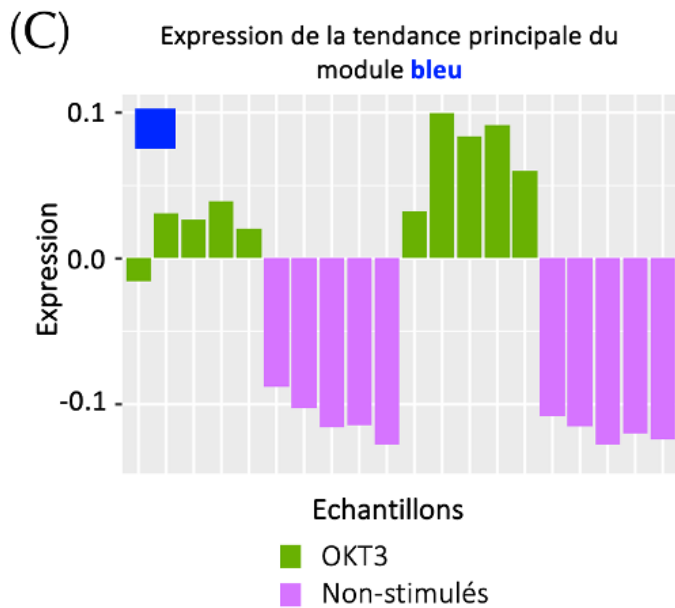


Figure 22: L'application de l'analyse WGCNA à un autre jeu de données disponible au laboratoire confirme les résultats déjà obtenus. Les échantillons de données transcriptomiques générés dans l'étude du rôle du facteur de transcription IRF4 dans l'épuisement ont été utilisés pour construire des réseaux de co-expression en suivant la même approche que décrite dans ce mémoire. **(A)** Un dendrogramme des mesures de co-expression entre chaque gène permet de visualiser les modules de gènes co-exprimés sous formes de branches. Chaque gène est associé ou non à un module par une légende de couleurs. Le module dans lequel se trouve IRF4 est marqué d'une flèche bleue. **(B)** Tableau des familles de gènes "Hallmarck" enrichis dans le module bleu (MsigDB). **(C)** Tendence principale du module bleu au sein de 20 échantillons de l'analyse. Les échantillons OKT3 ont été stimulés par des plaques de culture couvertes d'anticorps anti-CD3 (OKT3).

10. Perspectives

Il est important de rappeler que les lymphocytes T CD8⁺ infiltrant les tumeurs solides présentent différents niveaux d'expression de PD-1. Les lymphocytes T CD8⁺ ayant des niveaux intermédiaires d'expression de PD-1 (PD1^{int}) possèdent un phénotype de lymphocytes T activés et des fonctionnalités normales. Au contraire, les lymphocytes T CD8⁺ à forts niveaux d'expression de PD-1 (PD1^{high}) ont également un phénotype de lymphocytes T activés, mais avec une fonctionnalité réduite, ce qui est également caractérisé par une expression élevée de la protéine TOX. Toutes ces caractéristiques nous permettraient, aujourd'hui, d'isoler spécifiquement des lymphocytes T intra-tumoraux dysfonctionnels par cytométrie en flux. C'est notamment ce qui a été fait par l'équipe de Daniela Thommen en 2018 (Thommen, 2018). Ils ont isolés et comparés les transcriptomes de lymphocytes T CD8⁺ intra-tumoraux PD-1^{neg}, PD-1^{int} et PD-1^{high} infiltrant des NSCLC. Cette approche leur a permis d'identifier 8 clusters de gènes différentiellement exprimés et d'étudier les spécificités des transcriptomes des lymphocytes T CD8⁺ PD-1^{high}. Néanmoins, l'approche de génération des données utilisée dans cette étude ignore les lncRNAs ne possédant pas de queue de polyA. C'est donc la moitié des lncRNAs qui sont ignorés par cet article.

Notre laboratoire s'attelle actuellement à générer des données transcriptomiques de TILs CD8⁺ de cancers ovariens et de cancers pulmonaires en isolant les lymphocytes T CD8⁺ PD1^{int} et PD1^{high}. En utilisant le kit de construction de bibliothèques de séquençage Ovation SoLo (qui permet la détection des lncRNAs sans queue de polyA) et en utilisant les analyse développées dans ce mémoire, nous allons pouvoir réaliser une analyse transcriptomique pour identifier et étudier pour la première fois l'ensemble des lncRNAs liés à l'épuisement des lymphocytes T CD8⁺ PD1^{high}.

Matériels et Méthodes

1. Analyse bio-informatique

Toutes les analyses bio-informatiques ont été réalisées en utilisant le langage R via l'interface Rstudio (version 4.1.0). Le langage R est destiné aux statistiques ainsi qu'à l'analyse et la visualisation de jeux de données complexes. Le langage R dispose d'un grand nombre de *packages*. Un *package* est un ensemble de fonctions, de données et de documentation qui ajoutent de nouvelles fonctionnalités au langage de base. Une grande partie des packages utilisés dans ce travail font partie du projet Bioconductor (Gentleman et al., 2004 ; Huber et al., 2015) qui comprend plus de 2000 packages spécialement conçus pour l'analyse de jeux de données omiques.

1.1. Identification des gènes différentiellement exprimés

La normalisation du nombre de reads comptés sur chaque gène a été effectuée avec le package DESeq2 (Love, Huber et Anders, 2014). DESeq2 est un package qui permet d'identifier les gènes différentiellement exprimés en prenant en compte la profondeur de séquençage, ce qui n'est pas le cas dans d'autres analyses de ce type. Les gènes différentiellement exprimés entre nos échantillons d'ascites et nos échantillons sanguins ont été identifiés avec le même package. Pour ce test, un seuil de significativité fixé à 5% et des limites de *Fold-change* à 2 et 0,5. La représentation des données par PCA a été générée avec le package Factoextra (Kassambara et al., 2017). Les lncRNAs ont été identifiés à partir de la liste des lncRNAs annotés dans la base de données biologiques ENSEMBL (Cunningham et al., 2022).

1.2. Construction des réseaux de co-expressions.

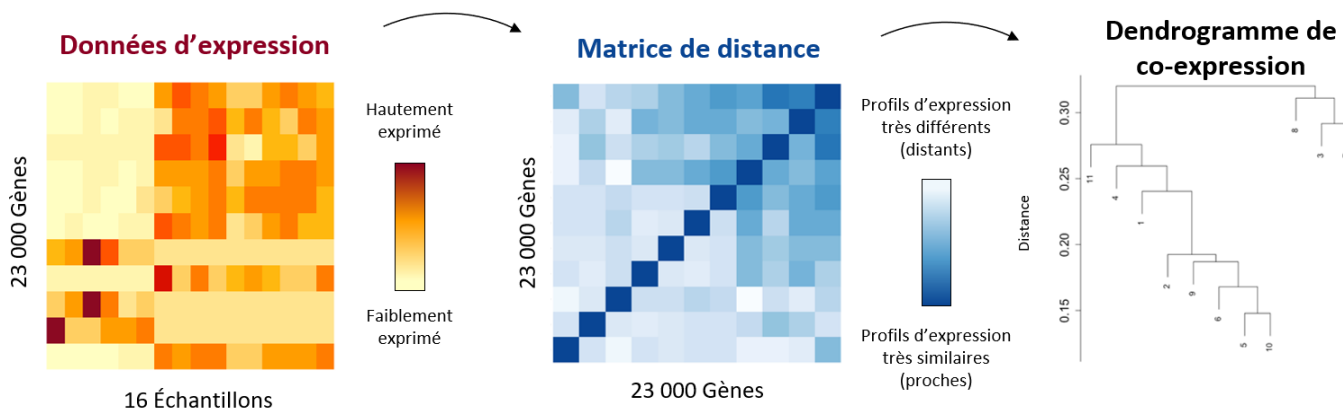


Figure 23 : Construction d'un dendrogramme de co-expression à partir des données d'expression génique.

La construction et l'analyse des réseaux de gènes co-exprimés a été menée avec le package WGCNA (Weighted Gene Co-Expression Network Analysis, version 1.71) (Langfelder et Horvath, 2008 et 2012). Ce package permet de construire des réseaux de co-expressions pondérés pour mieux comprendre les mécanismes biologiques sous-jacents à une maladie ou une condition, il permet également de découvrir de nouveaux gènes impliqués dans ces pathologies.

L'analyse par WGCNA se déroule comme ceci : une fois normalisées par DESeq2, les données d'expression sont utilisées pour générer une "matrice de distance" (*dissimilarity matrix*). Cette matrice contient les "distances" qui séparent chaque paire de gènes en fonction de leur similarité. Cette distance est déterminée par 2 facteurs : la corrélation de leur expression et leur chevauchement topologique (Topological Overlap). Le chevauchement topologique mesure à quel point deux gènes partagent des gènes co-exprimés dans un réseau. Dans la matrice de distance, moins la distance entre les gènes est importante, plus leur expression sera considérée comme similaire, et donc plus les gènes seront considérés comme co-exprimés.

Un clustering hiérarchique est ensuite appliqué à la matrice de distance. Le résultat de ce clustering est un graphique qui regroupe tous les gènes : Le dendrogramme de co-expression (**Figure 7, page 22**). Ce dendrogramme nous permet d'identifier, par un algorithme, les modules de gènes co-exprimés pour étudier les gènes codants qui sont co-exprimés à nos lncRNAs d'intérêt. (**Boîte 3 : Interpréter un dendrogramme de co-expression, page 22**).

1.3. Analyses d'enrichissement fonctionnel

Les modules ■ brun, ■ vert et ■ turquoise (**Figure 7, page 22**) ainsi que l'ensemble des gènes différentiellement exprimés ont été soumis à une analyse d'enrichissement fonctionnel par sur-représentation (*Over Representation Analysis*). Cette analyse compare le nombre de gènes d'un ensemble à des familles de gènes annotées pour leurs fonctions biologiques (aussi appelés ou *gene sets*).

Nous avons comparé nos ensembles d'intérêt à une collection de *gene sets* de la base de données MSigDB : la collection C7 (Subramanian, Tamayo, et al.,2005 ; Liberzon, et al.,2011). C'est plus précisément la sous-collection "Signalisation immunitaire" qui a été utilisée dans ce mémoire. Cette sous-collection contient plus de 4500 *gene sets* qui représentent une grande diversité de perturbations du système immunitaire, de types et d'états cellulaires.

2. Cultures cellulaires et milieux

Les lymphocytes T CD8⁺ isolés du sang de donneurs atteints d'hémochromatose ont été cultivés dans du milieu Iscove's Modified Dulbecco's Medium (IMDM) (21980065, Thermo Fisher) supplémenté avec 10% de sérum humain (HS) (préparation interne au laboratoire), 1,5mM GlutaMAX (35050061, ThermoFisher Scientific), 100U/ml pénicilline (P4333, Sigma-Aldrich), 100µg/ml streptomycine (P4333, Sigma-Aldrich), et de l'interleukine-2 (IL-2) (CLB-P-476, Proleukin, Novartis). Ce milieu est nommé "IMDM complet" dans la suite de ce mémoire. Les cellules ont été maintenues à 37°C et 8% de CO₂.

Les cellules HeLa ont été cultivées dans du milieu DMEM (31885023, Gibco™) supplémenté avec 10% de sérum foetal bovin (FBS) (F7524, Sigma-Aldrich), 4,5g/L de D-glucose (G8644, Sigma-Aldrich, 1mM de Sodium Pyruvate (113-24-6, JT baker), 100U/ml pénicilline et 100µg/ml streptomycine . Les cellules ont été maintenues à 37°C et 8% de CO₂.

3. Décongélation

Les cryotubes (377267, ThermoFisher Scientific) contenant les lymphocytes T CD8⁺ congelés ont été placés au bain marie à 37°C pendant 1 minute. Le contenu des

cryotubes a ensuite été transféré dans un Falcon de 15mL (525-0604, VWR European) avant l'ajout goutte à goutte de 10 mL de l'IMDM complet chaud le long du tube. Ces tubes ont été immédiatement centrifugés 7 minutes à 300g. Le surnageant a ensuite été aspiré et le culot a été resuspendu à $2 \cdot 10^6$ cellules/mL d'IMDM complet + 5U/mL de DNase I (7469, StemCell technologies) pour être réparti à 2,5mL par puits dans une plaque 24 puits (3526, Corning Incorporated). Ces plaques ont été incubées 18 heures à 37°C et 8% de CO₂.

4. Isolement des lymphocytes T CD8⁺ du sang de donneurs atteint d'hémochromatose

Les poches de sang ont été centrifugées pendant 20 minutes à 3300g et à 12°C (avec une décélération réduite). Le plasma a été retiré de la poche avant la récupération du buffy coat. Le buffy coat a ensuite été déposé sur une solution de Lymphoprep™ (1114547, Alere Technologies) puis centrifugé à 1000g pendant 20 minutes à accélération et décélération minimale. L'anneau de cellules mononuclées a été récupéré et lavé 3 fois dans du PBS-EDTA (1mM). La première centrifugation a été faite à 400g pendant 10 minutes, et les 2 suivantes à 300g pendant 7 minutes.

Les PBMCs ainsi obtenus ont ensuite été resuspendus à 10^7 cellules/mL dans du RPMI (s2400025, Gibco) et 1 mL de globules rouges de mouton par 10mL de PBMC (E-410, Virion Serion) y a été ajouté. Une fois que cette solution a été mélangée brièvement par inversion, elle a été déposée sur une solution de Lymphoprep™ (1114547, Alere Technologies) puis centrifugée à 1000g pendant 20 minutes à accélération et décélération minimale. L'entièreté du surnageant a ensuite été retiré et le culot contenant les globules rouges de moutons a été resuspendu dans 1,5mL de RPMI et 13.5 mL de solution de lyse (12mL de NH₄Cl et 1,5mL de FBS) afin de provoquer la lyse des globules rouges de mouton. Ensuite, les cellules ont été immédiatement centrifugées 1 minute à 2000rpm à RT. Le surnageant a ensuite été retiré et le culot a été resuspendu dans 50mL de RPMI.

Les PBMC CD2⁺ ont été comptées au bleu de trypan (15250061, Gibco™), centrifugées 7 minutes à 300g et resuspendues à 20^6 cellules/60μL dans du PBS-EDTA (1mM)-HS (1%) avec 20% de FcR block (130-059-901, Miltenyi) et 20% de microbilles CD8⁺ (130-045- 201,

Miltenyi) pendant 15 minutes à 4°C. Les cellules ont été passées sur un filtre de 40µm et triées par un AutoMACS (autoMACS® Pro Separator, Miltenyi) en mode POSSEL dans une solution de PBS-EDTA (1mM)-HS (1%). La fraction positive contenant les lymphocytes T CD8⁺ a été comptée au bleu de trypan. Les cellules ont été lavées avec de l'IMDM et congelées dans 50% de IMDM, 40% de HS et 10% de DMSO(Sc358801, Santa Cruz Biotech)

5. Marquages et cytométrie en flux

Les cellules ont été récoltées et transférées en plaque 96 puits à fond conique. La plaque a été centrifugée 4 minutes à 400g, 4°C. Les cellules ont été resuspendues dans 50µl de PBS-EDTA (1mM)-HS (1%) avant l'ajout de 50µL de PBS-EDTA (1mM)-HS (1%) contenant les anticorps appropriés. Après 20 minutes à 4°C, les cellules ont été lavées avec 100µl de PBS-EDTA (1mM)-HS (1%) et centrifugées 4 minutes à 400g, 4°C.

Les cellules ont été resuspendues dans du PBS-EDTA (1mM)-HS (1%) et analysées par cytométrie en flux avec le BD LSRII Fortessa (*BD Biosciences*). Les données ont été analysées sur le programme Flowjo (*BD Biosciences*).

Tableau 2 : Liste des anticorps utilisés pour les marquages

Cible	Fluorochrome	Clone	Dilution utilisée pour le marquage	Référence
CD8 β	PerCP Cy5.5	2ST8.5H7	1/100	Custom order, BD Bioscience
CD2	BV480	RPA-2.10	1/500	746538, BD Bioscience
β 2M	W632	W6/32	1/200	21HLAAAR, IMGEN
PD-1	BV786	EH12.2H7	1/40	329930, BD Bioscience
CD69	APC	FN50	1/75	310910, BD Bioscience
CD11a	PE	HI111	1/100	301207, BD Bioscience

6. Stimulation de lymphocytes T CD8

Les lymphocytes T CD8⁺ ont été cultivés en présence de billes couvertes d'anticorps anti-CD3 et anti-CD28 (11132D, Thermo Fisher) avec un ratio de 1 billes pour 10 lymphocytes T dans de l'IMDM complet avec de l'IL-2 (La concentration d'IL-2 utilisée dans chaque expérience est spécifiée dans les légendes des figures).

Lorsque les billes ont été retirées du milieu, la solution de cellules a été passée dans un aimant (12321D, DynaMag™) pendant 2 minutes, avant que le surnageant soit récupéré et incubé 48 heures supplémentaires à 37°C et 8% de CO₂

7. Nucléofection de siRNA

Une fois les cellules isolées et resuspendues dans 10mL de PBS-EDTA (1mM)-HS (1%), la solution de nucléofection P3 (V4XP-3024, Lonza) a été préchauffée à RT. Les mélanges de nucléofection ont été préparés sur glace en ajoutant 100 μ L de solution P3 à 5 μ L de siRNA à la concentration appropriée. Les cellules ont ensuite été centrifugées 7 minutes à 300g et à RT (Température ambiante) avant d'être resuspendues dans le mélange de nucléofection approprié. Ce mélange a été rapidement transféré dans une nucléocuvette (V4XP-3024, Lonza) pour être électroporées dans un 4D-Nucleofector® (AAF-1003X, Lonza) selon le programme "*T cell unstim, High efficiency*". Une fois cela fait, 1 mL de l'IMDM complet chaud a immédiatement été ajouté aux nucléocuvettes et le contenu de ces cuvettes a immédiatement été transféré dans les puits d'une plaque

48 puits. Une fois toutes les cellules nucléofectées et transférées dans la plaque, celle-ci est incubée 72 heures à 37°C et 8% de CO₂.

Tableau 3 : Liste des siRNA utilisés

Cible	5'-Séquence-3'	Référence	Nom du produit	Pureté
Négatif	Confidentiel	SIC005, Sigma-Aldrich.	MISSION® siRNA Fluorescent Universal Negative Control 1, Cyanine 5,	Standard
CD2	GAGAGGGUCUCAAACCAAtt	s225090, Thermofisher	Silencer® Select Pre-designed siRNA, Scientific.	Standard
B2M	UUAACUAUCUUGGGCUGUG	ab22432, abcam	Immunology Adaptive Immunity MHC Class I Share by email Anti-HLA Class I antibody [W6/32]	HPLC-RP

8. Traitement aux GapmeR

Une fois les lymphocytes T CD8⁺ isolés et centrifugés 7 minutes à 300g et 4°C, ils ont été resuspendus à 200 000 cellules/200µL dans de l'IMDM complet avec 10U/mL d'IL-2 et la concentration appropriée en GapmeR. Les 200µL de lymphocytes ont ensuite été transférés dans les puits d'une plaque 96 puits à fond rond et incubés 24, 48, ou 72 heures à 37°C et 8% de CO₂.

Tableau 4 : Liste des GapmeR utilisés

Cible	5'-Séquence-3'	Référence	Nom du produit
Négatif	6-FAM- AACACGTCTATACGC	339515, QIAGEN	LG00000002-DDB - Negative control A Antisense LNA GapmeR - 5' Fam
<i>CD11a 1</i>	GATGGTAGTGGCTGAG	339511, QIAGEN	LG00131461-DDA - CD11a_6
<i>CD11a 2</i>	ACGTCAATCATTAAAC	339511, QIAGEN	LG00168207-DDA - CD11a_12
<i>LINC00402</i>	CATGACATAACAGAAA	339511, QIAGEN	LG00196618-DDA - Human_linc00402_1

9. RT-qPCR

Les lymphocytes T CD8⁺ ont été resuspendus dans 700µL de Qiazol (79306, QIAGEN) et congelés dans l'azote liquide pour être stockés à -80°C. La solution a ensuite été décongelée avant d'être soumise à une extraction d'ARN sur colonnes par le kit miRNeasy Micro (217084, QIAGEN). Une fois extrait, l'ARN total de chaque échantillon a été ajouté à un mix de rétro-transcription (**Tableau 4**). La solution a été incubée 90 minutes à 42°C avant d'être refroidie sur glace. Un volume de 80µL d'eau distillée a alors été ajouté à la solution d'ADN complémentaire avant qu'elle soit stockée à -20°C.

Les échantillons d'ADN complémentaire ont ensuite été ajoutés au mix de qPCR préalablement préparé avec un des couples de primer dirigés contre EF-1, CD11a et LINC00402 selon ce qui est décrit dans le tableau 5. Le mélange a ensuite été incubé 30 secondes à 95°C avant d'entamer 40 cycles de PCR de 95°C pendant 10 secondes suivis de 60°C pendant 1 minute. La réaction a été conduite dans un QuantStudio™ 3 (A28567, *Applied Biosystems*)

Tableau 5 : Mix de rétrotranscription

	Concentration finale	Références
First-strand buffer 5x	1x	RevertAid Reverse transcriptase, EPO441, ThermoFisher Scientific
dNTP, 100mM	2mM	R1121, ThermoFischer Scientific
RNase inhibitor (Ribolock), 40U/µL	1U/µL	EO0381, ThermoFischer Scientific
DTT 100mM	10mM	y00147, Invitrogen
pd(N) ₆ primer 0.2µg/µL	5µM	s0142, ThermoFisher Scientific
M-MLV reverse transcriptase 200U/µL	10-5U/µL	RevertAid Reverse transcriptase, EPO441, ThermoFisher Scientific
ARN	1µg	

Tableau 6 : Mix de qPCR

	Concentration finale	Références
ROX SYBR® MasterMix Blue TTP	1x	UF-RSMT-B0701 Takyon™, EUROGENTEC
Primer sens	300 nM	Voir tableau 6
Primer anti-sens	300 nM	Voir tableau 6

Tableau 7 : Liste des primers utilisés pour les qPCR

Primers	Séquence	Références
EF1 : Primer sens	5'-GCTTCACTGCTCAGGTGAT-3'	Custom order, EUROGENTEC
EF1 : Primer anti-sens	5'-GCCGTGTGGCAATCCAAT-3'	Custom order, EUROGENTEC
CD11a : Primer sens	5'-ACCTGGTACATGTGCTTGAC-3'	Custom order, EUROGENTEC
CD11a : Primer anti-sens	5'-GACAACCTCAGCCACTACCATC-3'	Custom order, EUROGENTEC
LINC00402 : Primer sens	5'-GAAGTGGATATGGAAGCTTGGC-3'	Custom order, EUROGENTEC
LINC00402 : Primer anti-sens	5'-TTGGATGCCTCATGGGTTAGC-3'	Custom order, EUROGENTEC

10. Test de sécrétion de cytokines

Après les 72 heures d'incubation avec ou sans GapmeR, les lymphocytes T CD8⁺ ont été comptés au bleu de Türk (1-09277-0100, Merck). Les cellules ont ensuite été centrifugées 5 minutes à 400g avant d'être resuspendues à 10 000 cellules/100µL dans de l'IMDM complet et 10U/mL d'IL-2. Pour chaque puits d'une plaque 96 puits à fond conique (3894, Corning Incorporated), 100µL de cellules ont été ajoutées à 100µL de l'IMDM complet contenant 10U/mL d'IL-2 avec ou sans un total de 3333 billes αCD3/αCD28. La plaque a ensuite été centrifugée 1 minute à 400g avant d'être incubée 20h à 37°C et 8% de CO₂. Une fois cette incubation terminée, 150µL du surnageant de chaque puits ont été prélevés et transférés dans une plaque 96 puits à fond conique avant d'être scellés

et congelés à -80°C . Les surnageants respectifs ont ensuite été soumis à deux tests ELISA. Le test anti-TNF α a été réalisé suivant les protocoles fournis par le fabricant (430201, ELISA MAXTM Standard Set Human TNF- α , Biolegend).

Pour l'ELISA anti-IFN γ , une plaque 96 puits à fond plat "high-binding" (655061, Greiner Bio-one) a été recouverte de $4\mu\text{g}/\text{mL}$ d'un anticorps de capture anti-IFN γ (AHC4432, ThermoFischer Scientific) 18 heures à 4°C . La plaque a ensuite été lavée par 3 immersions dans un tampon de lavage (0,1% Tween 20 (P7949, Sigma-aldrich), 90g NaCl, H₂O) et une immersion dans de l'eau distillée. L'anticorps de détection (AHC4539, ThermoFischer Scientific) dilué dans de l'IMDM complet à une concentration de $0,5\text{mg}/\text{mL}$ a ensuite été transféré dans les puits de la plaque, $50\mu\text{L}$ par puits. Les échantillons et deux lignes de concentrations connues d'IFN γ (554616, BD Biosciences) ont ensuite été ajoutés, $50\mu\text{L}$ par puits, avant d'être incubés 2 heures à 37°C sous agitation. La plaque a ensuite été lavée une seconde fois de la même manière que décrite précédemment. Ensuite, $50\mu\text{L}$ de Streptavidine-HRP (N100, Biolegend) diluée dans du PBS + BSA 0,5% (A-7030, Sigma-Aldrich) ont été transférés dans chaque puits. La plaque a ensuite été incubée 30 minutes à 37°C en sous agitation avant d'être à nouveau lavée comme décrit précédemment. Un volume de $100\mu\text{L}$ de la solution TMB (34028, ThermoFisher Scientific) est déposé dans chaque puits. Pour arrêter la réaction, $25\mu\text{L}$ d' H_2SO_4 (12933634, Fisher Chemicals) ont ensuite été ajoutés dans chaque puits avant que la couleur des échantillons ne commence à être saturée. La plaque a ensuite été insérée dans un iMarkTM Microplate Absorbance Reader (1681135, Biorad) pour lire l'absorbance de chaque puits à 450 et 655 nm. Les données ont ensuite été analysées avec le logiciel Excel (Microsoft Corporation).

11. Test de prolifération

Après les 72 heures d'incubation avec ou sans GapmeR, les lymphocytes T CD8⁺ ont été comptés au bleu de Türk (1-09277-0100, Merck). Les cellules ont été resuspendues à $1.10^6/\text{ml}$ dans du PBS avec $5\mu\text{M}$ de Cell Trace Violet (CTV) (C34557, ThermoFisher Scientific) et incubées 20 minutes à 37°C et 8% de CO₂. Une partie des lymphocytes a été dans resuspendue dans du PBS pur sans CTV. Ensuite, 5 volumes de l'IMDM complet ont été ajoutés, puis après 5 minutes à température ambiante, les cellules ont été centrifugées 7 minutes à 300g. Les cellules ont été resuspendues à $3.10^5/\text{ml}$ dans de

l'IMDM complet avec 25U/ml d'IL-2. Par après, 100µLde cellules et 100µLde l'IMDM complet contenant 25U/mL d'IL-2 et 10 000 billes αCD3/αCD28 (11132D, ThermoFisher Scientific) ont été ajoutées par puits d'une plaque 96 puits à fond rond (3799, Corning Incorporated). Les cellules ont été centrifugées 1 minute à 400g avant d'être incubées pendant 4 jours à 37°C et 8% de CO₂. Au jour 4 après activation, les cellules ont été récoltées avant d'être marquées pour analyser par cytométrie en flux leur viabilité, l'expression de CD8β, PD-1, CD69, et la dilution du marqueur CTV.

Bibliographie

Abewe H, Deshmukh S, Mukim A, Beliakova-Bethell N. Use of GapmeRs for gene expression knockdowns in human primary resting CD4+ T cells. *J Immunol Methods*. janv 2020;476:112674.

Ahmed R, Salmi A, Butler LD, Chiller JM, Oldstone MB. Selection of genetic variants of lymphocytic choriomeningitis virus in spleens of persistently infected mice. Role in suppression of cytotoxic T lymphocyte response and viral persistence. *Journal of Experimental Medicine*. 1 août 1984;160(2):521-40.

Albert R. Scale-free networks in cell biology. *Journal of Cell Science*. 1 nov 2005;118(21):4947-57.

Atianand MK, Caffrey DR, Fitzgerald KA. Immunobiology of Long Noncoding RNAs. *Annu Rev Immunol*. 26 avr 2017;35(1):177-98.

Barrena MJ, Echaniz P, Garcia-Serrano C, Zubillaga P, Cuadrado E. Differential expression of lymphocyte function-associated antigen (LFA-1) on peripheral blood leucocytes from individuals with Down's syndrome. *Clin Exp Immunol*. avr 1992;88(1):41-4.

Cabili MN, Dunagin MC, McClanahan PD, Biaisch A, Padovan-Merhar O, Regev A, et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biology*. 29 janv 2015;16(1):20.

Capece T. The Role of Integrin LFA-1 in CD8+ T cell Activation and Cell Fate. 2016 [cité 11 janv 2023]; Disponible sur: <https://urresearch.rochester.edu/institutionalPublicationPublicView.action?institutionalItemVersionId=31062>

Capece T, Walling BL, Lim K, Kim KD, Bae S, Chung HL, et al. A novel intracellular pool of LFA-1 is critical for asymmetric CD8+ T cell activation and differentiation. *J Cell Biol*. 6 nov 2017;216(11):3817-29.

Chen YG, Satpathy AT, Chang HY. Gene regulation in the immune system by long noncoding RNAs. *Nat Immunol*. sept 2017;18(9):962-72.

Cheng S, Li F, Qin H, Ping Y, Zhao Q, Gao Q, et al. Long Noncoding RNA lncNDEPD1 Regulates PD-1 Expression via miR-3619-5p in CD8 + T Cells. *J Immunol*. 15 mars 2022;208(6):1483-92.

Coulie PG, Boon T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *cancer*. 2014;12.

Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. *Nucleic Acids Research*. 7 janv 2022;50(D1):D988-95.

Dehmani S, Nerrière-Daguin V, Néel M, Elain-Duret N, Heslan JM, Belarif L, et al. SIRPy-CD47 Interaction Positively Regulates the Activation of Human T Cells in Situation of Chronic Stimulation. *Front Immunol*. 2021;12:732530.

Demotte N, Wieërs G, Van Der Smissen P, Moser M, Schmidt C, Thielemans K, et al. A galectin-3 ligand corrects the impaired function of human CD4 and CD8 tumor-infiltrating lymphocytes and favors tumor rejection in mice. *Cancer Res*. 1 oct 2010;70(19):7476-88.

Doering TA, Crawford A, Angelosanto JM, Paley MA, Ziegler CG, Wherry EJ. Network analysis reveals centrally connected genes and pathways involved in CD8+ T cell exhaustion versus memory. *Immunity*. 14 déc 2012;37(6):1130-44.

Fain JS, Lorient A, Diacofotaki A, Van Tongelen A, De Smet C. Transcriptional overlap links DNA hypomethylation with DNA hypermethylation at adjacent promoters in cancer. *Sci Rep*. 30 août 2021;11:17346.

Fazil MHUT, Ong ST, Chalasani MLS, Low JH, Kizhakeyil A, Mamidi A, et al. GapmeR cellular internalization by macropinocytosis induces sequence-specific gene silencing in human primary T-cells. *Sci Rep*. 24 nov 2016;6:37721.

Froehlich J, Versapuech M, Megrelis L, Largeteau Q, Meunier S, Tanchot C, et al. FAM65B controls the proliferation of transformed and primary T cells. *Oncotarget*. 27 sept 2016;7(39):63215-25.

- Gallimore A, Glithero A, Godkin A, Tissot AC, Plückthun A, Elliott T, et al. Induction and exhaustion of lymphocytic choriomeningitis virus-specific cytotoxic T lymphocytes visualized using soluble tetrameric major histocompatibility complex class I-peptide complexes. *J Exp Med.* 4 mai 1998;187(9):1383-93.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology.* 15 sept 2004;5(10):R80.
- Goyal A, Myacheva K, Groß M, Klingenberg M, Duran Arqué B, Diederichs S. Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. *Nucleic Acids Res.* 17 févr 2017;45(3):e12.
- Halaby MJ, Hezaveh K, Lamorte S, Ciudad MT, Kloetgen A, MacLeod BL, et al. GCN2 drives macrophage and MDSC function and immunosuppression in the tumor microenvironment. *Sci Immunol.* 13 déc 2019;4(42):eaax8189.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 4 mars 2011;144(5):646-74.
- Huang D, Chen J, Yang L, Ouyang Q, Li J, Lao L, et al. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *Nat Immunol.* oct 2018;19(10):1112-25.
- Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods.* févr 2015;12(2):115-21.
- Khan O, Giles JR, McDonald S, Manne S, Ngjow SF, Patel KP, et al. TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. *Nature.* juill 2019;571(7764):211-8.
- Kong Y, Hsieh CH, Alonso LC. ANRIL: A lncRNA at the CDKN2A/B Locus With Roles in Cancer and Metabolic Disease. *Front Endocrinol (Lausanne).* 24 juill 2018;9:405.
- Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell.* 25 janv 2018;172(3):393-407.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 29 déc 2008;9(1):559.
- Lemieszek MB, Findlay SD, Siegers GM. CellTrace™ Violet Flow Cytometric Assay to Assess Cell Proliferation. New York, NY: Springer US; 2022 [cité 11 janv 2023]. p. 101-14. (Methods in Molecular Biology). Disponible sur: https://doi.org/10.1007/978-1-0716-2376-3_9
- Li M, Zheng J, Luo D, Xu K, Zhang X. Tiki proteins are substrates of membrane-type matrix metalloproteinases. *FEBS Lett.* juill 2022;596(14):1851-9.
- Li Y, Syed J, Sugiyama H. RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem Biol.* 17 nov 2016;23(11):1325-33.
- Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 15 juin 2011;27(12):1739-40.
- Liu SJ, Nowakowski TJ, Pollen AA, Lui JH, Horlbeck MA, Attenello FJ, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* déc 2016;17(1):67.

- Lorenzi L, Chiu HS, Avila Cobos F, Gross S, Volders PJ, Cannoodt R, et al. The RNA Atlas expands the catalog of human non-coding RNAs. *Nat Biotechnol.* nov 2021;39(11):1453-65.
- Mantei A, Rutz S, Janke M, Kirchhoff D, Jung U, Patzel V, et al. siRNA stabilization prolongs gene knockdown in primary T lymphocytes. *European Journal of Immunology.* 2008;38(9):2616-25.
- Mantei A, Rutz S, Janke M, Kirchhoff D, Jung U, Patzel V, et al. siRNA stabilization prolongs gene knockdown in primary T lymphocytes. *European Journal of Immunology.* 2008;38(9):2616-25.
- Marsland BJ, Kopf M. T-cell fate and function: PKC- θ and beyond. *Trends Immunol.* avr 2008;29(4):179-85.
- Mondal T, Subhash S, Vaid R, Enroth S, Uday S, Reinius B, et al. MEG3 long noncoding RNA regulates the TGF- β pathway genes through formation of RNA-DNA triplex structures. *Nat Commun.* 24 juill 2015;6:7743.
- Nakano M, Ito M, Tanaka R, Yamaguchi K, Ariyama H, Mitsugi K, et al. PD-1+ TIM-3+ T cells in malignant ascites predict prognosis of gastrointestinal cancer. *Cancer Science.* 2018;109(9):2986-92.
- Nakano M, Ito M, Tanaka R, Yamaguchi K, Ariyama H, Mitsugi K, et al. PD-1+ TIM-3+ T cells in malignant ascites predict prognosis of gastrointestinal cancer. *Cancer Sci.* sept 2018;109(9):2986-92.
- Nie J, Zhao Q. Lnc-ITSN1-2, Derived From RNA Sequencing, Correlates With Increased Disease Risk, Activity and Promotes CD4+ T Cell Activation, Proliferation and Th1/Th17 Cell Differentiation by Serving as a ceRNA for IL-23R via Sponging miR-125a in Inflammatory Bowel Disease. *Front Immunol.* 2020;11:852.
- Nixon BG, Li MO. Satb1: Restraining PD1 and T Cell Exhaustion. *Immunity.* 17 janv 2017;46(1):3-5.
- Ntanasis-Stathopoulos I, Fotiou D, Terpos E. CCL3 Signaling in the Tumor Microenvironment. *Adv Exp Med Biol.* 2020;1231:13-21.
- Paley MA, Kroy DC, Odorizzi PM, Johnnidis JB, Dolfi DV, Barnett BE, et al. Progenitor and terminal subsets of CD8+ T cells cooperate to contain chronic viral infection. *Science.* 30 nov 2012;338(6111):1220-5.
- Park EG, Pyo SJ, Cui Y, Yoon SH, Nam JW. Tumor immune microenvironment lncRNAs. :25.
- Peltier D, Radosevich M, Ravikumar V, Pitchiaya S, Decoville T, Wood SC, et al. RNA-seq of human T cells after hematopoietic stem cell transplantation identifies Linc00402 as a regulator of T cell alloimmunity. *Sci Transl Med.* 17 mars 2021;13(585):eaaz0316.
- Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet.* janv 2016;17(1):47-62.
- Rashidi A, Miska J, Lee-Chang C, Kanojia D, Panek WK, Lopez-Rosas A, et al. GCN2 is Essential for CD8+ T-cell Survival and Function in Murine Models of Malignant Glioma. *Cancer Immunol Immunother.* janv 2020;69(1):81-94.

- Ravindran R, Loebbermann J, Nakaya HI, Khan N, Ma H, Gama L, et al. The amino acid sensor GCN2 controls gut inflammation by inhibiting inflammasome activation. *Nature*. 24 mars 2016;531(7595):523-7.
- Scott AC, Dündar F, Zumbo P, Chandran SS, Klebanoff CA, Shakiba M, et al. TOX is a critical regulator of tumour-specific T cell differentiation. *Nature*. juill 2019;571(7764):270-4.
- Scott-Browne JP, López-Moyado IF, Trifari S, Wong V, Chavez L, Rao A, et al. Dynamic Changes in Chromatin Accessibility Occur in CD8+ T Cells Responding to Viral Infection. *Immunity*. 20 déc 2016;45(6):1327-40.
- Shademan M, Naseri Salanghuch A, Zare K, Zahedi M, Foroughi MA, Akhavan Rezayat K, et al. Expression profile analysis of two antisense lncRNAs to improve prognosis prediction of colorectal adenocarcinoma. *Cancer Cell Int*. 2019;19:278.
- Song Z, Wu W, Chen M, Cheng W, Yu J, Fang J, et al. Long Noncoding RNA ANRIL Supports Proliferation of Adult T-Cell Leukemia Cells through Cooperation with EZH2. *J Virol*. 15 déc 2018;92(24):e00909-18.
- Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics*. 2016;54(1):1.30.1-1.30.33.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 25 oct 2005;102(43):15545-50.
- Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun*. 17 oct 2019;10(1):4706.
- Thommen DS, Koelzer VH, Herzig P, Roller A, Trefny M, Dimeloe S, et al. A transcriptionally and functionally distinct PD-1+ CD8+ T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat Med*. juill 2018;24(7):994-1004.
- Thommen DS, Schreiner J, Müller P, Herzig P, Roller A, Belousov A, et al. Progression of Lung Cancer Is Associated with Increased Dysfunction of T Cells Defined by Coexpression of Multiple Inhibitory Receptors. *Cancer Immunol Res*. déc 2015;3(12):1344-55.
- Thommen DS, Schumacher TN. T Cell Dysfunction in Cancer. *Cancer Cell*. avr 2018;33(4):547-62.
- Tili E, Michaille JJ, Cimino A, Costinean S, Dumitru CD, Adair B, et al. Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-alpha stimulation and their possible roles in regulating the response to endotoxin shock. *J Immunol*. 15 oct 2007;179(8):5082-9.

van Dam S, Vösa U, van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. *Briefings in Bioinformatics*. 20 juill 2018;19(4):575-92.

Vermot Desroches C, Rigal D. Leukocyte function-associated antigen-1 expression on peripheral blood mononuclear cell subsets in HIV-1 seropositive patients. *Clinical Immunology and Immunopathology*. 1 août 1990;56(2):159-68.

Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. :18.

Webb J, Milne K, Kroeger D, Nelson B. PD-L1 expression is associated with tumor-infiltrating T cells and favorable prognosis in high-grade serous ovarian cancer. *Gynecologic oncology*. 10 mars 2016;141.

Wherry EJ, Ha SJ, Kaech SM, Haining WN, Sarkar S, Kalia V, et al. Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity*. oct 2007;27(4):670-84.

Ye M, Xie M, Zhu J, Wang C, Zhou R, Li X. LPS-Inducible lncRNA TMC3-AS1 Negatively Regulates the Expression of IL-10. *Front Immunol*. 22 juill 2020;11:1418.

Yin J, Ma Z, Selliah N, Shivers DK, Cron RQ, Finkel TH. Effective gene suppression using small interfering RNA in hard-to-transfect human T cells. *Journal of Immunological Methods*. 30 mai 2006;312(1):1-11.

PCA - Principal Component Analysis Essentials - Articles - STHDA [Internet]. [cité 11 janv 2023]. Disponible sur: <http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/112-pca-principal-component-analysis-essentials/>

Annexes

Annexe 1 : Paramètres de l'algorithme de construction des modules d'expression génique

Paramètre	Valeur
Type de réseau	Non-signés
Puissance de Soft-Thresholding	9
Taille minimale des modules	15 gènes
Sensibilité de l'algorithme (Entre 0 et 4)	1
Fusion des modules	Non
Étape de PAM	Non

Abréviations

aCD3/aCD28	couvert d'anticorps anti-CD3 et anti-CD28
aB2M	Anticorps anti- 2M
aCD11a	Anticorps anti-CD11a
aCD2	Anticorps anti-CD2
aCD8b	Anticorps anti-CD8
ceRNA	ARN compétitif endogène
ELISA	Méthode immuno-enzymatique ELISA
GVHD	Maladie du greffon contre l'hôte
lncRNA	Long ARN non-codant
NSCLC	Cancer bronchique non à petite cellules
RT	Température ambiante
siRNA	Petit ARN interférant
TCR	Récepteur des lymphocytes T
WGCNA	Analyse des réseaux pondérés de co-expression génique.

The immune system is able to recognize and eliminate cancer cells. Nonetheless, when T lymphocytes are chronically activated by cancer cells, they gradually lose their effector functions. This phenomenon is usually referred to as "exhaustion" and could explain the proliferation of the cancer cells that escape immune control. Understanding this phenomenon seems to be a relevant approach to improve immunotherapy treatments, such as T-cell transfer therapies.

The role of long non-coding RNAs (lncRNAs) in exhausted T cells has received very little attention, particularly in humans. Previously considered as useless genomic noise, many lncRNAs are now known to play essential roles, particularly in CD8⁺ T cell effector functions.

The first objective of my thesis was to develop a bioinformatic approach to identify lncRNAs potentially involved in the exhaustion of intratumoral CD8⁺ T lymphocytes. We developed an approach combining the analysis of the differentially expressed lncRNAs between two conditions with the analysis of co-expressed gene networks. This analysis allows the identification of lncRNAs related to a phenomenon of interest while characterizing their potential functions by analyzing their known co-expressed coding genes. At the end of this analysis, a dozen lncRNAs potentially involved in exhaustion were identified. Among them, LINC00402 was chosen for in-depth characterization.

The second objective of my thesis was to develop a functional characterization method for these lncRNAs. The electroporation of small interfering RNAs was tested in CD8⁺ T lymphocytes and proved disappointing. We used GapmeRs. Preliminary experiments showed that the use of a GapmeR targeting LINC00402 has an effect on proliferation and the production of IFN γ and TNF α of CD8⁺ T lymphocytes.