

République Algérienne Démocratique et populaire

Ministère de L'enseignement Supérieur et De la Recherche Scientifique

Université de Blida 01



Faculté Des Science De La Nature Et De La Vie

Département de Biologie et Physiologie cellulaire BPC

Mémoire de fin d'étude en vue de l'obtention du diplôme de master en

Spécialité : **Génétique**

Filière : **Sciences biologique**

Domaine : **Sciences de la Nature et de la Vie**

Intitulé :

Etude « in silico » et identification de quelques oncogènes (& variants) grâce au modèle tumoral, cas des leucémies.

Présenté par

- ✓ **Nasserine AIB**
- ✓ **Zineb BENDEKFAL**

Devant le jury composé de :

Présidente et Examinatrice :	Dr Wahiba Hamzi	MCA	Université Blida 01
Promoteur :	Dr Mohamed Said Ramdane	MCA	Université Blida 01
Co-Promotrice:	Dr Simona Abdulhussein Alia	MCA	Université Blida 01

Année Universitaire : 2020 – 2021.

Remerciements

Nous remercions Dieu, le tout puissant de nous avoir donné la santé, la volonté, le courage d'achever ce mémoire de fin d'études avec beaucoup de passion, et de beaux souvenirs de moments précieux, de travail d'équipe en binôme et avec la bienveillance de nos chers Professeurs et encadreurs.

Ces quelques lignes ne suffiront pas à exprimer toute notre gratitude et chaleureux remerciements à notre promoteur Dr Med Said Ramdane & notre CoPromotrice Dr Simona Abdulhoussein Alia, Maitres de conférences à l'université de Blida 1, d'avoir accepté de nous encadrer pour ce projet de mémoire. Nous leurs témoignons notre sincère reconnaissance pour leurs disponibilité, encouragements et grande gentillesse, nous avons adoré travailler avec eux.

Au Dr Wahiba Hamzi, Maitre de conférences à l'université de Blida 1, de nous avoir fait l'honneur d'examiner notre mémoire et de présider le jury, merci pour toute la pertinence et la générosité de ses remarques et de sa bienveillance.

Ce mémoire nous a permis de rencontrer humainement des personnes merveilleuses et d'une rare générosité, nos Professeurs, camarades et administration de notre faculté, nous tenons à les remercier chaleureusement, merci Mr Aichi.

A tous ceux qui nous ont inspiré, ainsi qu'à ceux qui ont contribué de près ou de loin à la réalisation de notre présent mémoire.

Merci à toutes et tous.

Dédicaces

Mon modeste travail, en BioMédical, réalisé en guise d'immersion pratique de fin d'études, est orienté vers le développement de la vie et du bien être de l'être humain à travers l'ingénierie génomique.

*Cela, en vue de développer des protocoles biomédicaux de soins de médecine prédictive ciblée personnalisée intégrative systémique génétique, des soins par les thérapies douces, la thérapie génique et les biothérapies biotechnologiques, pour **soulager la douleur et donner des chances de vie et de bien être aux patients ayant les profils génétiques concernés**, raison de ma reprise d'études, dont je fais une mission, de ce qui reste de mon humble vie.*

*J'ai adoré réaliser mon rêve, malgré tout ce que j'ai vécu en cette période, celui **d'éditer la vie, les gènes de vie**, à 37 ans et de débloquer les actions dans ce sens.*

*Je dédie ce travail donc, en premier lieu, à tous ceux qui nous faciliteront notre mission divine et humaine de scientifiques entrepreneurs. Je le dédie à vous : **Ceux que j'aime**.*

À tous ceux qui ont été bons, respectueux et généreux humainement avec moi, m'ont donné leur confiance, et mes chances méritées, À tous ceux qui ont cru et investi en moi et en mes engagements à un moment donné de ma vie. À ceux qui ont marqué mon âme par leur excellence humaine, leur discipline, engagements, et leurs valeurs nobles.

A Melodia , Yakoota , Théo et biens d'autres , Je ferai de mon mieux pour réunir les plus sensibles à vos douleurs, les meilleurs, les plus engagés , promis ☺

À ma tendre et douce Algérie qui m'accompagne partout. MERCI.

À ma grande famille de cœur à travers le monde, ces gens qui vivent par passion pour honorer leur mission divine, les âmes athlètes, Nobles. À l'âme de tous ceux que j'aime partis trop tôt. Reposez en paix.

À mes amis, mes amours, mes collègues, camarades, à mes douces et si passionnantes et passionnées Beenzu & Zineb, à ceux qui m'aident à faire du bon travail, et à mes compagnons du voyage, celui de la vie Sereine Papillon & bon enfant ☺ Merci de me rendre Heureuse en votre compagnie inestimable, à tous nos moments sincères.

À l'Afrique ☺ Mon cœur t'appartient

*Ma sincère gratitude va aux Dr Mohammed said R, Dr Errahmani B , Dr S Alia et à tous mes Professeurs, Enseignants et Encadreurs. Que Dieu vous ouvre toutes ses portes ici bas et au Paradis, jamais j'oublierai ce que vous avez fait pour moi, à un moment aussi difficile dans ma Vie. Dr Mohamed Belhocine, merci cher ami, **Merci biens chers**.*

Merci à toutes et tous de m'inspirer.

***Bon travail et bonne chance** à nous tous, je suis à votre écoute, n'hésitez pas à me contacter.
#Imagine #GiveLifeChances*

Nasserine Serine

Email : Serine.biology@gmail.com

Dédicaces

Je dédie ce travail à mes parents, qui peuvent trouver ici le résultat de longues années de sacrifices, souffrances et de privations pour m'aider à avancer dans la vie.

Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de vous ; je vous aime.

Une spéciale dédicace pour :

Mes chères soeurs Meriem et Chaima et Mes frères Abd alaziz et Abd errahmen et Houdaifa source de joie et de bonheur

Mes oncles : Nourdine, Hocine et Mohammed

Mes sincères copines : Sarah Meddah ,Sara Boulifa , Amina Bouras, Assia ,Jihad ,warda Hiba ,Asma et Ratiba.

Pour tous les membres de ma promotion

A tous mes enseignants depuis mes premières années d'études.

Toute ma gratitude va à mes chers professeurs durant mon cursus universitaire, mais une mention très Spéciale va au Dr Mohammed said .

Et sans oublier mon adorable amie, ma sœur, ma belle, ma binôme Nasserine Aib et sa famille, je lui souhaite tout le bien, le bonheur, et beaucoup de succès dans sa vie personnelle et professionnelle.

A tous ceux qui m'ont aidée et soutenu pendant tout mon cursus universitaire. Que dieu les les protège et leur réalise leurs rêves et souhaits.

Enfin je tiens à dédicacer ce travail à ma mère et ma tante Sabah parties trop tôt, mais dont je sais qu'elles auraient été fières de moi.

Zineb

Liste des abréviations

ADN : Acide désoxyribonucléique

ARN : Acide ribonucléique ou

NGS : Séquençage de nouvelle génération

NCBI: Le National Center for Biotechnology Information

NIH : The nation's medical research agency

Liste des figures

Figure 1 : L'Hématopoïèse	07
Figure 2 : Les altérations génétiques	08
Figure 3 : Les étapes de la cancérogenèse.....	09
Figure 4 : Les atteintes génotoxiques	10
Figures 5 : Les mécanismes génétiques des tumeurs	15
Figure 6 : Le génome	24
Figure 7 : Protocol d'étude in silico	25
Figure 8 : Protocol d'étude in silico de l'oncogène MYC	31
Figure 9 : Protocol d'étude in silico de l'oncogène HOX11.....	35
Figure 10 : Protocol d'étude in silico de l'oncogène LMO2	38
Figure 11 : Protocol d'étude in silico de l'oncogène EVI1.....	40
Figure 12 : Protocol d'étude in silico de l'oncogène LAZ3	42

Liste des tableaux

Tableau 1 : Listing des oncogènes et leurs variantes leucémiques

Résumé

Les leucémies sont des maladies clonales et acquises de la cellule souche hématopoïétique ou d'un précurseur déjà commis vers les lignées lymphoïde et/ou myéloïde.

Le clonage moléculaire de remaniements chromosomiques récurrents a permis d'identifier de nombreuses altérations géniques qui constituent des outils diagnostiques et de suivi thérapeutique maintenant pris en compte dans les nouveaux critères de classification des leucémies.

Les domaines de certains gènes marquent les oncogènes clés et sont associés à des gènes dérégulés dans le cancer, dans notre thèse – Les leucémies.

De vastes domaines de gènes ont été associés après leur séquençage à des oncogènes leucémiques.

Notre objectif est d'identifier les oncogènes leucémiques et leurs variants *in silico* en nous basant sur des recherches documentaires sur des travaux d'études d'identification *in vivo* déjà réalisés sur les oncogènes leucémiques, en vue d'initier une approche de diagnostic prédictif de cas pareils de leucémies par séquençage génomique, basé sur des données d'études réalisées *in vivo* disponibles sur les plateformes bioinformatiques.

Ici, nous avons identifié les gènes associés aux domaines leucémies, les oncogènes et leurs variants.

Nous étudierons théoriquement l'expression de gènes dérégulés dans les leucémies, à savoir les variants oncogéniques afin de prouver que les leucémies humaines est un processus de transformation par étapes selon les mutations des oncogènes et de leurs variants.

Ceci par l'utilisation de données open-source du séquençage de différents modèles et d'un tas de pipelines bioinformatiques.

Le développement récent de thérapeutiques ciblées sur l'activité ou la stabilité d'une protéine oncogénique a été illustré ici, par les succès thérapeutiques obtenus dans le traitement des leucémies.

Mots clés : Leucémie, oncogènes, variantes, mutations, bioinformatique.

Summary

Leukemias are clonal and acquired diseases of the hematopoietic stem cell or a precursor already committed to the lymphoid and / or myeloid lineages.

The molecular cloning of recurrent chromosomal rearrangements has made it possible to identify numerous gene alterations which constitute diagnostic and therapeutic monitoring tools now taken into account in the new criteria for classifying leukemia.

The domains of certain genes mark key oncogenes and are associated with deregulated genes in cancer, in our thesis - Leukemias. Large domains of genes have been associated after sequencing with leukemic oncogenes.

Our objective is to identify leukemic oncogenes and their variants in silico based on documentary research on work on in vivo identification studies already carried out on leukemic oncogenes, with a view to initiating a predictive diagnostic approach of similar cases of leukemia by genomic sequencing, based on data from studies carried out in vivo available on bioinformatic platforms.

Here, we have identified genes associated with leukemia domains, oncogenes and their variants. We will theoretically study the expression of deregulated genes in leukemias, namely oncogenic variants in order to prove that human leukemia is a process of transformation in stages according to the mutations of oncogenes and their variants.

This is through the use of open-source data from the sequencing of different models and a bunch of bioinformatic pipelines.

The recent development of therapies targeting the activity or stability of an oncogenic protein has been illustrated here by the therapeutic successes obtained in the treatment of leukemia.

Keywords: Leukemia, oncogenes, variants, mutations, bioinformatics.

ملخص

اللوكيميا هي مرض نسيلىة ومكتسبة من الخلايا الجذعية المكونة للدم أو سلائف ملتزمة بالفعل بالأنساب اللمفاوية و / أو النخاعية.

أتاح الاستنساخ الجزيئي لعمليات إعادة ترتيب الكروموسومات المتكررة إمكانية تحديد العديد من التعديلات الجينية التي تشكل أدوات مراقبة تشخيصية وعلاجية تؤخذ الآن في الاعتبار في المعايير الجديدة لتصنيف سرطان الدم.

تحدد مجالات جينات معينة الجينات المسرطنة الرئيسية وترتبط بجينات غير منظمة في السرطان ، في أطروحتنا - اللوكيميا.

ارتبطت مجالات كبيرة من الجينات بعد التسلسل مع الجينات المسرطنة اللوكيميا.

يتمثل هدفنا في تحديد الجينات المسرطنة لسرطان الدم ومتغيراتها في السيليكو بناءً على بحث وثائقي عن العمل في دراسات تحديد الهوية في الجسم الحي التي تم إجراؤها بالفعل على الجينات الورمية لسرطان الدم ، بهدف بدء نهج تشخيصي تنبؤي لحالات مماثلة من سرطان الدم عن طريق التسلسل الجيني ، بناءً على البيانات من الدراسات التي أجريت في الجسم الحي متاحة على منصات المعلومات الحيوية.

هنا ، حددنا الجينات المرتبطة بمجالات سرطان الدم والجينات الورمية ومتغيراتها.

سوف ندرس نظرياً التعبير عن الجينات المحررة في اللوكيميا ، أي المتغيرات الورمية من أجل إثبات أن سرطان الدم البشري هو عملية تحول على مراحل وفقاً لطفرات الجينات المسرطنة ومتغيراتها.

يتم ذلك من خلال استخدام بيانات مفتوحة المصدر من تسلسل نماذج مختلفة ومجموعة من خطوط أنابيب المعلوماتية الحيوية.

تم توضيح التطور الأخير في العلاجات التي تستهدف نشاط أو استقرار البروتين المسبب لسرطان هنا ، من خلال النجاحات العلاجية التي تم الحصول عليها في علاج سرطان الدم.

الكلمات المفتاحية: اللوكيميا ، الجينات المسرطنة ، المتغيرات ، الطفرات ، المعلوماتية الحيوية.

Sommaire

CHAPITRE 1 : Rappels bibliographiques

1.1 Les leucémies	06
1.1. a Définition & Classification	06
1.1. b La leucémogénèse - Cancérisation	06
1.1. c Atteintes génotoxiques - épigénétiques	09
1.1. d Instabilité génétique	11
1.2 Les Oncogènes	14
1.2. a Notions de base : Les oncogènes – Proto Oncogènes - Oncoprotéines	14
1.2. b Oncogénèse	15
1.2. c Mutations oncogéniques : Variants des oncogènes	16
1.3 Les Biothérapies	16
1.3. a Analyse génomique : In vivo (Cas cancer acquis) - In Silico de modélisation et prédiction	16
1.3. b Tests génétiques – Séquençage – Profil génétique – Carte génétique	17
1.3. c Etudes in vivo épidémiologiques	18
1.3. d'Etudes in silico et biobanques	18
1.3. e Thérapie génique ciblée personnalisée	18
1.3. f Les anti-oncogènes : gènes suppresseurs de tumeurs	19
1.3. g Biomarqueurs leucémiques pronostiques et prédictifs	20

CHAPITRE 2 : Partie Pratique expérimentale In Silico

2.1 Matériel	22
- Matériel Non biologique	22
2.2 Méthodes	22
- Principe d'étude : Diagnostic de Prédiction préventif & Modélisation in silico	22

- Protocol d'étude : Stratégie d'identification in silico	22
- Recherche documentaire (Bibliographie d'études in vivo)	23
- Analyse & Modélisation BioInformatique : Etapes d'identification in silico	23
• Interrogation des bases de données & génération des données à haut débit	23
Outils : NCBI – NIH- Blast Fasta	23

CHAPITRE 3 : Résultats et Discussions

3.1 Résultats :	27
- Résultats Bibliographiques	27
- Résultats bioinformatiques	27
Cas pratiques d'identification d'oncogènes et variants leucémiques in silico : Recherche documentaire et in silico.	40
3.2 Analyse & Discussion	43
<u>Conclusion et perspectives</u>	52
<u>Références bibliographiques</u>	56
<u>Annexes</u>	59

Glossaire

La thérapie génique ou **généthérapie** est une stratégie thérapeutique qui consiste à faire pénétrer des gènes dans les cellules ou les tissus d'un individu pour traiter une maladie. La thérapie génique vise à remplacer ou compléter un allèle mutant défectueux par un allèle fonctionnel ou à surexprimer une protéine dont l'activité aurait un impact thérapeutique.

Médecine personnalisée : La **médecine de précision**, également appelée **médecine personnalisée**, a pour objectif de proposer au patient un traitement adapté aux caractéristiques de sa tumeur.

Génomique : La **génomique** est la science des **génomés** : elle étudie les séquences d'ADN des êtres vivants. Un **génome** est formé de l'ensemble des informations génétiques contenues dans la cellule.

Effecteurs épigénétiques : La caractérisation d'inhibiteurs spécifiques de certains *effecteurs épigénétiques* a ouvert une nouvelle voie thérapeutique, la thérapie épigénétique,

La biologie systémique : Une **biologie intégrative de données globalisées**, c'est l'étude des interactions entre les composants d'un système biologique.

Signatures génomiques : Une **signature génomique** correspond à une sélection de gènes permettant de mieux appréhender le pronostic ou de mieux prédire l'efficacité d'un traitement.

Génétique prédictive en oncologie médicale : La médecine prédictive désigne les capacités nouvelles de la médecine, et notamment de la génétique de prévoir, parfois très longtemps à l'avance, les affections qui frapperont le patient. La prédiction basée sur le code génétique est limitée par l'interaction complexe des gènes, des facteurs environnementaux et l'histoire biologique des individus.

Biopuce ou Puce à ADN/Protéine : Une puce à ADN est un ensemble de molécules d'ADN fixées en rangées ordonnées sur une petite surface qui peut être du verre, du silicium ou du plastique.

Oncogène : Les oncogènes sont une catégorie de gènes dont l'expression favorise la survenue de cancers. Ce sont des gènes qui commandent la synthèse d'oncoprotéines, protéines stimulant la division cellulaire ou inhibant la mort cellulaire programmée, ce qui déclenche une prolifération désordonnée des cellules.

Introduction

Contexte de l'étude : Le décryptage du génome humain, au début du XXI^e siècle, a facilité l'analyse du fonctionnement cellulaire sous l'influence des gènes. D'où l'entrée en scène d'une génomique fonctionnelle cancérologique qui s'attache notamment à comprendre, afin de les contrôler, les mécanismes de l'oncogenèse qui sont à l'origine de la tumorigenèse. Elle conduit ainsi à une rénovation de la cancérologie, en permettant d'agir sur des cibles très précises, précocement identifiées, et en augmentant les possibilités thérapeutiques grâce à une définition plus rigoureuse des lieux à traiter.

Les nouveaux patients de cancer entretiennent de l'espoir quant à l'évolution des connaissances scientifiques associées au génome humain et au développement de technologies et de médicaments prometteurs, susceptibles de leur offrir une meilleure chance de vaincre le cancer ou de prolonger leur vie et d'en améliorer la qualité.

La majorité des cancers résulte d'évènements mutationnels de l'ADN somatique et plus rarement d'anomalies constitutionnelles héréditaires. La détection d'anomalies génétiques au sein des cellules cancéreuses à l'aide de différentes technologies a mis en évidence des biomarqueurs dont la signification clinique est connue ou émergente. Le schéma conceptuel qui sous-tend ainsi notre compréhension des bases fondamentales de la tumorigenèse implique l'intervention, le plus souvent aléatoire, de gènes clés qui contrôlent des processus cellulaires essentiels, comme la prolifération et la croissance des cellules, leur mort programmée (apoptose), leur sénescence, les communications inter- et intracellulaires, la motilité cellulaire, et la formation de vaisseaux (angiogenèse). La compréhension de ces processus physiopathologiques des leucémies a été marquée par la découverte, au milieu des années 1970, des premiers gènes déclencheurs de tumeurs (oncogènes), puis, environ une dizaine d'années plus tard, par celle des gènes suppresseurs de tumeurs ; antioncogènes.

De fait, l'essor des technologies génomiques et des thérapies ciblées permet, entre autres en oncologie, de repérer les gènes de susceptibilité aux cancers, de préciser un diagnostic ou un pronostic et d'optimiser la pharmacothérapie selon le profil génétique d'un individu. Par ailleurs, l'approche expérimentale *in vivo* consiste à rechercher plusieurs mutations ciblables à l'aide du séquençage d'un grand panel de gènes, de l'exome ou du génome, et à traiter le patient avec une molécule hors indication reconnue ou en essai clinique.

Le séquençage des génomes fournit, en effet, aux expérimentateurs un nombre considérable de séquences qui peuvent être utilisées comme sondes génétiques pour les études et diagnostics *in silico*. Les signatures génomiques et les informations génétiques contenues dans la tumeur permettent de mieux traiter les patients et parfois de ne pas les traiter lorsque les caractéristiques prédisent un bon pronostic. Par contre, les résultats préliminaires d'études des profils moléculaires *in vitro* obtenus en cours sont prometteurs et à terme, de nouvelles données guideront la prise en charge clinique des cancers Séquençage génétique des cancers. Il est très probable avec la collecte des données génétiques à grande échelle en cours et le développement parallèle des thérapies ciblées que des patients de plus en plus nombreux bénéficient de cette médecine prédictive et personnalisée. **Guylaine Rouleau ,2015.**

La Révolution numérique d'analyse moléculaire & prédiction des oncogènes des leucémies *in silico*, à savoir la bio-informatique est une discipline émergente de la recherche qui se place à l'interface de la biologie et de l'informatique.

Par opposition aux classiques manipulations *in vivo* ou *in vitro* pratiquées en laboratoire, on parlera alors d'expériences « *in silico* » (néologisme d'allure semi-latine, formé à partir de l'anglais silicon). Ce domaine de l'analyse de données biologiques par ordinateur a de plus en plus tendance à se dénommer « bioanalyse ». **Erwin van Dijk, 2021.**

La thérapie génique, comme son nom l'indique, consiste notamment à aller réparer le gène défaillant au cœur de la cellule, qui peut ainsi accomplir à nouveau son travail correctement. L'ennui, c'est que c'est encore plus compliqué que pour une maladie génétique, notamment dans le cas des leucémies. Outre le fait qu'il faut trouver un vecteur inoffensif pour faire rentrer le gène sain dans le noyau de la cellule, le problème avec les cellules cancéreuses leucémiques, c'est qu'il y a plusieurs gènes défectueux à l'intérieur. Et pour qu'elle ne soit plus cancéreuse, il faudrait tous les réparer.

Les leucémies font aujourd'hui l'objet d'une analyse génétique dite somatique, à la recherche d'une thérapie ciblée efficace.

Objectif

La présente étude donne une idée sur des notions générales d'identification *in silico* qui portent sur les anomalies génétiques des leucémies , le séquençage intégral et le séquençage ciblé de ces tumeurs précisément à l'aide des technologies de séquençage de nouvelle génération (NGS) *in*

silico et les interventions théranostiques, notamment les diagnostics prédictifs in silico de cas cas mutationnels leucémiques semblables.

Problématique

Notre problématique est de répondre à cet objectif de diagnostic prédictif et de traitement génique ciblé des cas leucémiques, il revient à nous donc aujourd'hui dans ce mémoire d'identifier cette multitude de gènes défectueux, d'oncogènes leucémiques ainsi que leurs variants, et les étudier in silico à travers la recherche bibliographique des études in vivo et des données haut débit via des outils bioinformatiques.

Résultats attendus

Quelques résolutions génomiques à partir de l'étude des variants d'oncogènes identifiés dans le cas des leucémies, en vue de réaliser un diagnostic prédictif dans des cas pareils et d'identifier une thérapie génique ciblée pour ces **profils mutationnels de leucémies**.

CHAPITRE 1

Rappels
Bibliographiques

1.1 Les leucémies

1.1.a Définition & Classification :

Une orchestration précise qui se dérègle.

Schématiquement, on distingue les leucémies en fonction de deux critères principaux :

- Leur évolution: on différencie ainsi les leucémies chroniques, qui se développent progressivement, et les leucémies aiguës qui se déclarent plus brutalement ;

- La nature des cellules anormales : on parle de leucémies myéloïdes lorsque ces cellules sont issues des cellules souches myéloïdes, et de leucémies lymphoïdes lorsqu'elles se sont formées à partir des cellules souches lymphoïdes.

(Sylvie Gisselbrecht,2003)

Epidémiologie des leucémies

Les différents types de leucémies aiguës représentent entre 10 et 15 % des hémopathies malignes et sont des cancers rares selon la définition donnée par le groupe européen RARECARE avec un taux d'incidence annuel standardisé à la **population mondiale inférieur à 6 cas pour 100 000 habitants.**

1.1.b La leucémogénèse – Cancérisation

- **La cellule souche hématopoïétique et l'hématopoïèse:**

Le système hématopoïétique est organisé de façon hiérarchique avec au sommet les CSH (0.01 à 0.005 %) qui possèdent quatre propriétés fondamentales :

- L'auto renouvellement : capacité de se multiplier indéfiniment sans subir de différenciation. Cette propriété permet de maintenir un pool de cellules souches.

- La capacité de prolifération

- La multi potentialité: capacité à se différencier et à donner naissance à toutes les cellules de l'hématopoïèse.

- La capacité de différenciation (ARGAGY, 2021)

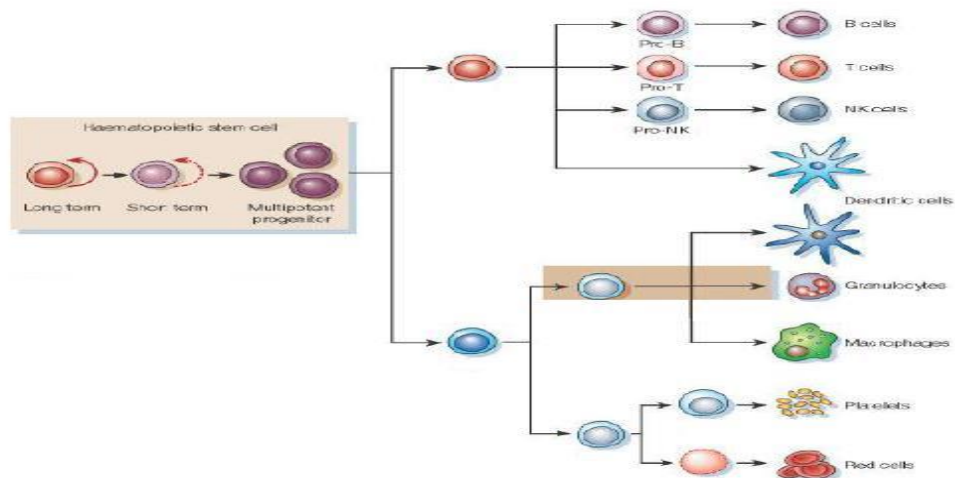


Figure 1 : L'Hématopoïèse

Etude du profil mutationnel des leucémies aiguës myéloïdes du sujet âgé par séquençage haut débit : résultats de l'essai ALFA-0701

(Sylvie Gisselbrecht, 2003)

- **La leucémogénèse :**

Certaines mutations confèrent un avantage prolifératif à la cellule : elles sont 39 appelées « driver », d'autres au contraire ne possédant pas cet effet sont dites « passager » [48]. Cependant, cette classification binaire est dépendante également du contexte .

La notion de cellules souches leucémiques (CSL) a été mise en évidence en 1994 [45].

Pour assurer sa transformation en CSL, la cellule hématopoïétique normale doit acquérir (ou conserver) des propriétés d'auto-renouvellement, voir ses propriétés de réponse aux signaux prolifératifs et antiprolifératifs modifiées, acquérir une résistance à l'apoptose et perdre ses capacités de différenciation. Il s'agit d'un phénomène multi-étape qui résulte de l'accumulation de multiples anomalies génétiques acquises : des translocations équilibrées, des mutations somatiques, une altération du nombre de copie de gènes mais également une hétérogénéité de l'expression génique [46] et du profil de méthylation de l'ADN .

- **Le cancer – Les leucémies : Maladie du gène - Mutations génétiques :**

Récemment, les mutations fréquemment rencontrées dans les leucémies ont été classées en huit catégories fonctionnelles principales. (ARGAGY, 2021)

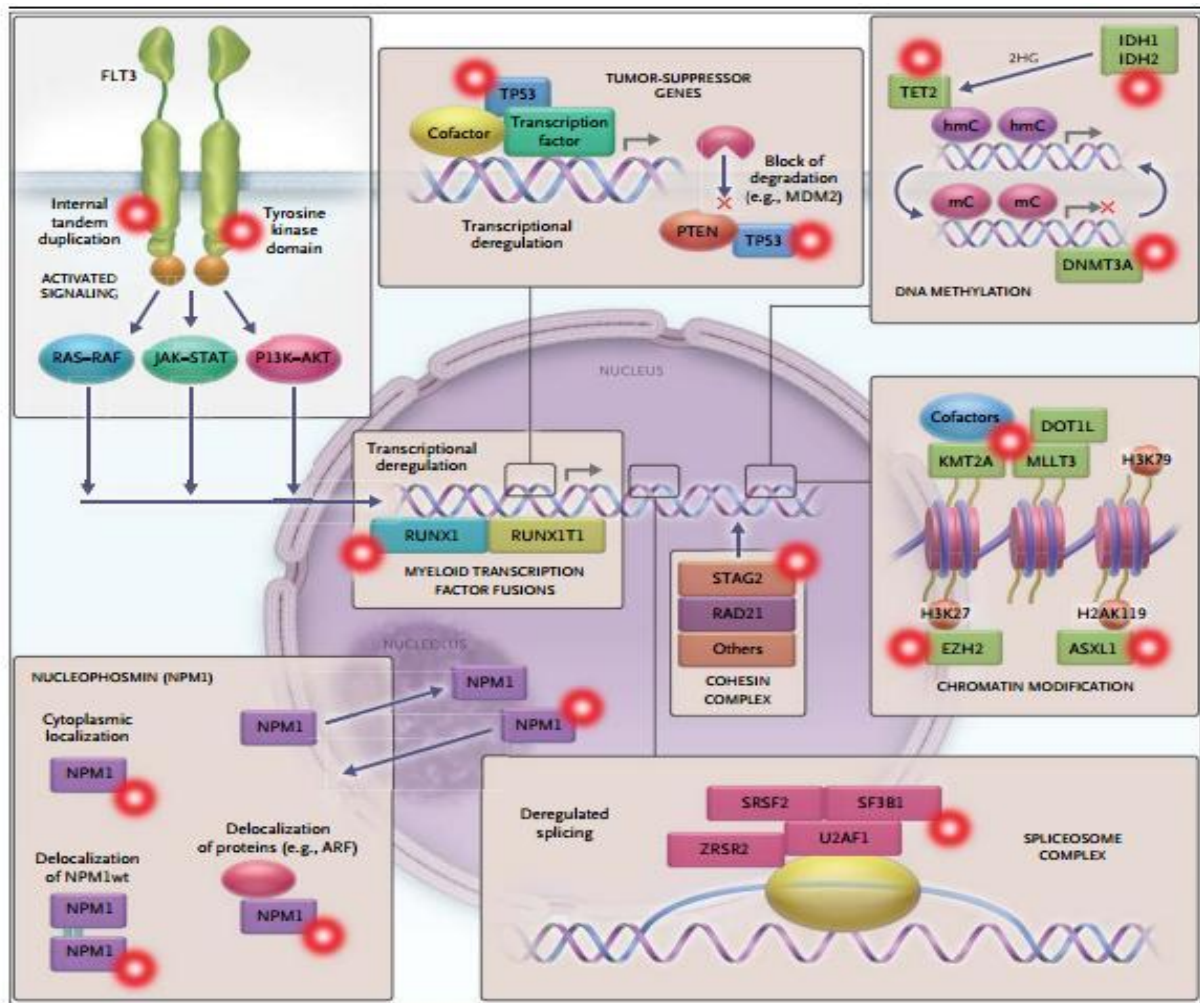


Figure 2 : Les altérations génétiques

Représentation schématique des 8 classes d'altérations génétiques responsables du phénomène de leucémogénèse. Ref : Etude du profil mutationnel des leucémies aiguës myéloïdes du sujet âgé par séquençage haut débit : résultats de l'essai ALFA-0701 (Sylvie Gisselbrecht, 2003)

- Cancérisation :

Le mécanisme essentiel de la carcinogénèse repose sur l'activation de certains oncogènes (gènes intervenant dans le contrôle de la prolifération cellulaire et capables de provoquer un processus de cancérisation) et sur l'inactivation d'antioncogènes (gènes suppresseurs de cancers, contrôlant l'intégrité du génome).

- Une ou plusieurs mutations

Les chercheurs ont tendance à diviser les cancers en deux grands groupes principaux en fonction du nombre d'altérations génétiques.

a. Les tumeurs pauci-génétiques (*pauci = peu*)

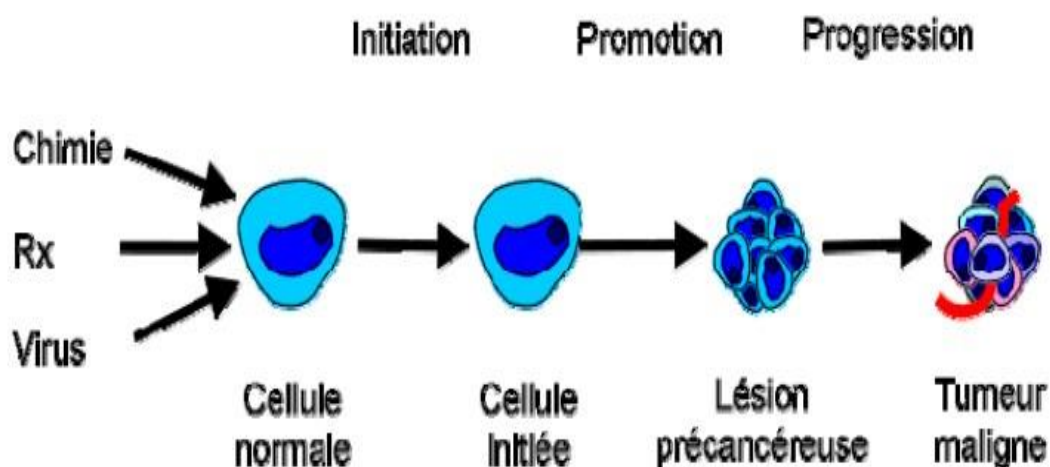
Ces tumeurs présentent une ou très peu d'anomalies génétiques. Les accidents génétiques apparaissent tôt dans la vie.

b. Les tumeurs à oncogénèse complexe

Ce sont les plus fréquentes. Les anomalies observées concernent tous les aspects du vivant :

- Les mécanismes hérités du vivant unicellulaire
- Le contrôle du cycle cellulaire
- La réparation de l'ADN
- La résistance aux xénobiotiques
- Les mécanismes acquis par les organismes plus élaborés
- Les processus d'intégration de l'information
- La réponse aux hormones et aux facteurs de croissance
- La réponse adaptative aux modifications du micro-environnement
- Les communications intercellulaires
- L'acquisition de fonctions biologiques particulières..., etc. (ARGAGY, 2021)

1.1.C Atteintes génotoxiques - épigénétiques



*Figure 3 : Les étapes de la cancérogenèse
(Charlotte Bouard, 2016)*

- **1^{ère} étape : Initiation**

Atteinte de l'ADN par un cancérigène génotoxique dit « initiateur », c'est un phénomène irréversible.

a. *Cellules initiées* : échappent au contrôle normal de division cellulaire

Agents initiateurs (génotoxiques) :

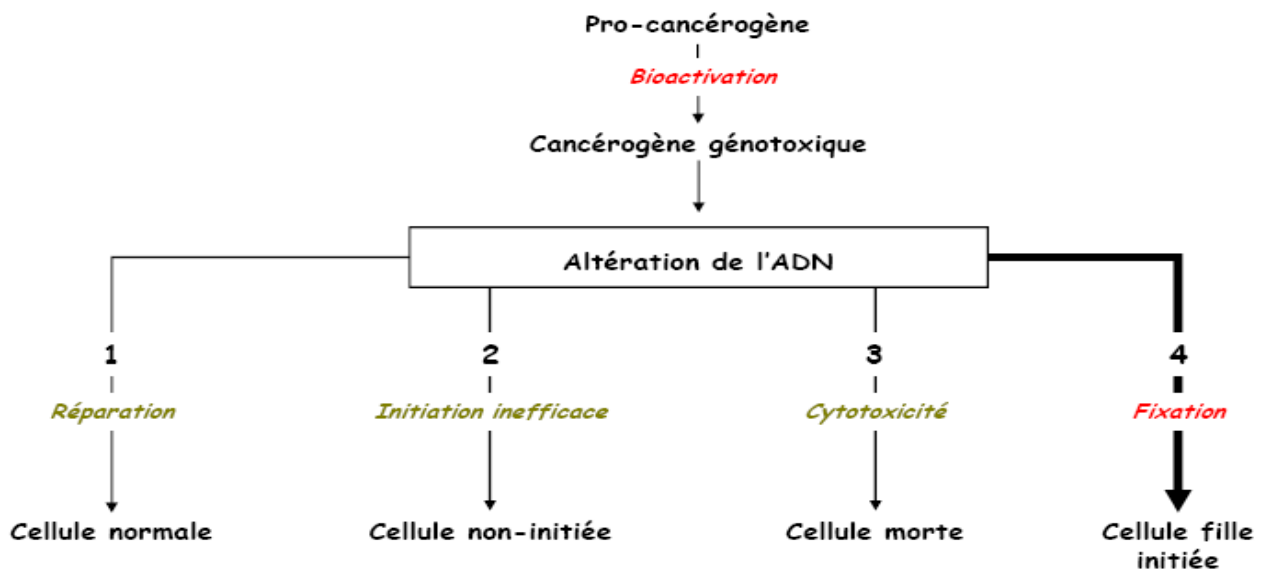
- agents chimiques (les plus nombreux)
- agents biologiques (virus ...)
- agents physiques (radiations ionisantes, UV ...)

b. *Les atteintes génotoxiques :*

- Mutations géniques : changement des séquences d'ADN des gènes (substitution, insertion ou délétion)
- Mutations chromosomiques : changement de la structure du chromosome (délétion, duplication, inversion ou translocation)

Mutations génomiques : changement du nombre de chromosome (aneuploidie, polyploidie).
(Charlotte Bouard, 2016)

Figure 4 : Les atteintes génotoxiques :



(Charlotte Bouard, 2016)

- **2^{ème} étape : Promotion** Processus épigénétique (non génotoxique) induit par un promoteur entraînant la stimulation de la sélection des cellules initiées. C'est un phénomène réversible.

Mécanismes :

- Sélection positive : stimulation directe des cellules initiées (promoteurs mitogènes).
- Sélection négative : destruction des cellules normales (promoteurs cytotoxiques) et meilleure croissance des cellules initiées.

- **3ème étape : Progression**

- Phase clinique : étape finale dans le développement du cancer.
- La néoplasie est cliniquement détectable
- Survenue de métastases.

1.1.d Instabilité génétique

Au fur et à mesure du développement de la maladie, les cellules tumorales acquièrent une instabilité génétique de plus en plus marquée, aboutissant à des mutations de plus en plus nombreuses.

La majorité des cancers chez l'homme se caractérisent par une instabilité chromosomique aussi appelée CIN (*chromosomal instability*).

La CIN est un phénomène dynamique se traduisant par le gain ou la perte de chromosomes entiers ainsi que par des remaniements structuraux des chromosomes. Ce processus contribue donc à l'aneuploïdie (cellule qui ne possède pas le nombre normal de chromosomes) des cellules malignes et à la constitution d'une hétérogénéité intratumorale.

Tout au long de la vie, les cellules se divisent pour se reproduire à l'identique et remplacer les cellules qui meurent. Chaque fois que la cellule se divise, elle doit recopier la totalité des 3 milliards de bases A, C, G et T. (**Nicholas J Short ans all 2018**)

Les cellules peuvent faire des erreurs en recopiant la chaîne ADN. Il peut s'agir, par exemple :

- De mutations : A T GGC devenant A A GGC
- De délétions : AT GGCC TGG devenant après la perte de GGCC , ATTGG
- D'insertions : ATGGC devenant après insertion de AAA , AT AAA GGC
- De translocations, par exemple entre le chromosome 9 et le chromosome 22, t (9;22)

- **La modélisation**

Elle est basée sur l'hypothèse de Knudson dites " two hits " développée pour le rétinoblastome. Elle a permis de modéliser la séquence des événements qui aboutit à la maladie.

- *Un premier événement aboutit à une mutation d'un gène.*

Cette première mutation peut être :

- Germinale , dans les formes de cancers familiaux
- Somatique , dans les formes sporadiques de la maladie

Cette mutation conduit à l'hétérozygotie. A ce stade, un seul des deux chromosomes est le siège d'une mutation, d'où la formule CO.

- *Un second événement entraîne une mutation somatique.* Elle aboutit à la perte de l'hétérozygotie. Maintenant, les deux chromosomes présentent une mutation, d'où la formule CC. Cette seconde mutation permet à la maladie de se développer.

- *A l'échelle moléculaire :*

On considère actuellement qu'un cancer est la résultante d'une ou plusieurs des anomalies suivantes:

- *Une activation des oncogènes qui codent pour la production de protéines contrôlant la prolifération cellulaire, l'apoptose ou les deux,*
- *Une inhibition des gènes suppresseurs de tumeurs*
- *Une inhibition des gènes encodant des micro-ARN, impliqués dans la régulation de certains gènes.*

- *En pratique :*

- *La production d'une protéine anormale*
- *L'augmentation de la production d'une protéine normale.* InfoCancer Arcagy.

Il faut distinguer les tests qui apportent des informations sur le patrimoine génétique transmissible, présent dans toutes les cellules de l'organisme (**génétique constitutionnelle**), et les tests qui informent sur l'état du génome de cellules tumorales (**génétique somatique**). En outre, d'autres tests permettent d'obtenir des informations sur la réponse à un traitement ou sur les risques d'effets secondaires (**pharmacogénomique**). (Nicholas J Short ans all 2018)

- *Les tests de génétique somatique : Séquençage génétique*

Les tests de génétique somatique (non héréditaire) consistent à **analyser le génome des cellules cancéreuses pour détecter des mutations survenues spécifiquement dans la tumeur et prédire la réponse à un traitement ciblé.** On parle de « **test compagnon** ».

Deux principales approches d'analyses du génome sont utilisées dans le cadre des tests génétiques :

- La **génétique moléculaire**, qui consiste à analyser la molécule d'ADN pour **détecter des mutations ou autres anomalies de façon ciblée sur le génome.** Elle fait appel à des techniques de biologie moléculaire.

- La **cytogénétique**, qui consiste à étudier le nombre et la forme des chromosomes pour **détecter des remaniements affectant des fragments chromosomiques ou des chromosomes entiers**. Le plus souvent, cette étude s'appuie sur l'observation du caryotype du patient, correspondant à la photographie de l'ensemble de ses chromosomes. C'est l'analyse la plus répandue chez les fœtus, avec près de 41 000 caryotypes réalisés en 2012.

Ces deux approches sont de plus en plus perméables, notamment de part l'utilisation croissante d'outils communs tels que **les puces à ADN**.

Les puces à ADN permettent de **détecter des remaniements chromosomiques cent fois plus petits que ne le permettait jusque-là le caryotype**. Dès lors, le risque devient celui de **découvrir une autre maladie que celle recherchée** ... Le comité d'éthique de l'Inserm se penche actuellement sur cette question : que faire d'une information importante pour la santé découverte de façon fortuite, que l'individu n'a pas demandé à savoir et n'est pas forcément désireux d'entendre ?

Ceci nous amène à l'importance du **Conseil génétique** et des **consultations oncogénétiques prédictives** par conséquence.

Grâce à la **génétique moléculaire** qui permet de préciser le **statut génétique** d'une personne pour une pathologie donnée, le conseil ne se limite plus à un simple calcul mathématique basé sur les données familiales, les modes de transmission, les données recueillies lors d'enquêtes.

Mieux informé sur la maladie, la hauteur réelle du risque encouru, la possibilité ou non de pouvoir recourir à un diagnostic prénatal, le couple à risque peut ainsi décider d'avoir ou non l'enfant qu'il désire. (**ARGAGY, 2021**)

1.2. Les Oncogènes

1.2. a Notions de base : Les oncogènes – Proto Oncogènes - Oncoprotéines

Proto Oncogènes : Tout gène cellulaire, appelé proto-oncogène (c-onc), susceptible de devenir, par suite d'une modification qualitative ou quantitative, un gène transformant, c'est-à-dire un gène capable de conférer expérimentalement le phénotype cancéreux (transformation) à une cellule normale eucaryote.

L'altération d'un allèle est suffisante pour entraîner une activation anormale.

Les oncogènes ont été classiquement définis comme des gènes, à caractère dominant, qui, lorsqu'ils sont exprimés de façon dérégulée ou lorsque leur structure est altérée, contribuent au phénotype transformé d'une cellule. Ces gènes codent pour des protéines de fonctions très diverses.

Les oncogènes sont répartis en 6 grandes classes en fonction des oncoprotéines pour lesquels ils codent :

1. les facteurs de croissance (assurent une boucle de régulation autocrine),
Exemple : proto-oncogènes codant pour les protéines de la famille FGF (fibroblast growth factor)
2. les récepteurs transmembranaires de facteurs de croissance
Exemple : le proto-oncogène erb B code pour le récepteur à l'EGF (epidermal growth factor)
3. les G-protéines ou protéines membranaires liant le GTP
Exemple : proto-oncogènes de la famille ras
4. les tyrosines protéine-kinases membranaires
5. les protéine-kinases cytosoliques
6. les protéines à activité nucléaire : contrôlent la transcription de gènes cibles en interagissant avec l'ADN.

Les anomalies les plus fréquentes sont principalement des inversions intrachromosomiques ou des translocations qui résultent de l'échange réciproque de matériel génétique entre deux chromosomes.

Ces altérations sont acquises et présentes dans toutes les cellules leucémiques d'un même patient, ce qui témoigne de l'origine clonale de la leucémie.

Leur caractère souvent récurrent, voire leur association systématique à certains types d'hémopathies suggéraient qu'elles pouvaient être l'événement causal du processus de transformation. **Il semblait donc possible que des oncogènes connus ou à découvrir étaient localisés à proximité des points de cassure chromosomique.**

1.2. b Oncogénèse

- *Activation des oncogènes : Carcinogénèse*

L'étude moléculaire des oncogènes a permis de découvrir plusieurs types d'activation pouvant correspondre à plusieurs étapes de la carcinogénèse. La survenue d'une ou de plusieurs de ces activations déclenche un cancer. La carcinogénèse chimique correspond à une activation par mutation dans un oncogène. Elle se caractérise par deux stades successifs : l'initiation et la promotion.

L'amplification d'un oncogène correspond à une activation par multiplication du nombre de copies de l'oncogène dans la cellule.

- *Altérations génétiques en l'absence d'anomalie cytogénétique*

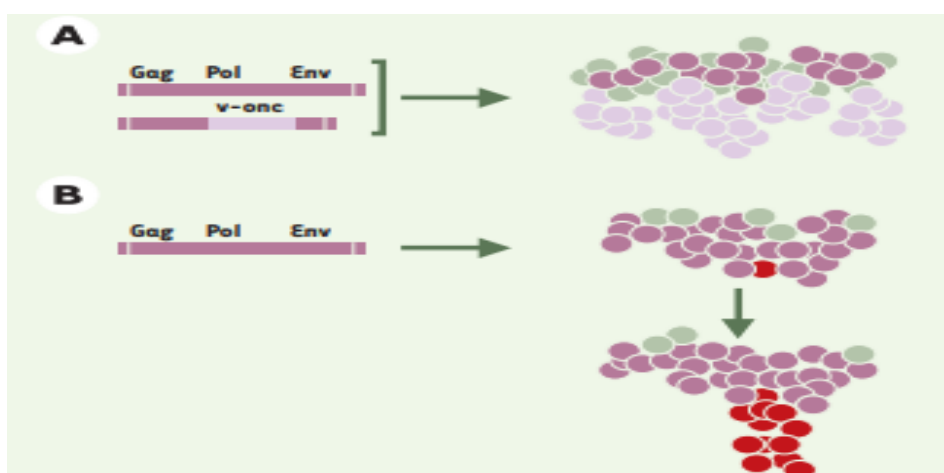
Des mutations ponctuelles et des microdélétions non détectables par les méthodes les plus performantes de cytogénétique, peuvent aussi conduire à l'activation d'oncogènes.

Ref : Oncogènes et leucémies: historique et perspectives Oncogenes in leukemia: historical background and therapeutic prospect. Sylvie Gisselbrecht

Il s'agit souvent d'un événement tardif au cours de la progression tumorale.

L'amplification de l'A.D.N. ou de l'A.R.N. d'un oncogène conduit à l'augmentation d'une protéine pouvant lutter contre le produit d'un autre gène, un gène suppresseur de la transformation tumorale (antioncogène). La quantité de protéine oncogène l'emporterait sur celle de la protéine suppresseur de tumeur, provoquant ainsi le cancer.

D'autres facteurs sont impliqués : activité anormale de gènes contrôlant des enzymes (télomérases) et de gènes bloquant l'apoptose et permettant la formation de nouveaux vaisseaux sanguins qui alimentent le tissu tumoral. (Marlon Arnone and All, 2020)



Figures 5 : Les mécanismes génétiques des tumeurs

Représentation schématique des mécanismes à l'origine des tumeurs viro-induites. A (ARGAGY, 2021)

1.2. c Mutations oncogéniques : Variants des oncogènes

- *Les facteurs d'oncogénèse sont :*

- **Intégration virale :** La carcinogenèse virale comprend deux sortes d'activation. La première implique des virus lents qui, par insertion, activent un oncogène. La seconde concerne des virus qui portent un gène transformant, capable de transformer la cellule en cellule tumorale.
- **Mutation ponctuelle :** Dans une séquence codante pour un proto-oncogène aboutissant à une modification fonctionnelle de l'oncoprotéine. Les mutations faux-sens entraînant la substitution d'un acide aminé par un autre, sont capables d'activer des proto-oncogènes en oncogènes, en touchant par exemple un site catalytique ou en entraînant une activation substitutive de la protéine.
- **Délétion :** Les délétions, qui aboutissent le plus souvent à une perte de fonction, peuvent parfois entraîner une activation anormale si elles touchent une région régulatrice.
- **Réarrangement structural**
Des altérations chromosomiques (translocations, inversions...) peuvent avoir pour conséquence moléculaire la formation d'un gène hybride généré par la fusion de régions codantes entraînant la synthèse de protéines chimériques non fonctionnelles.
- **Amplification génique**
L'amplification correspond à une augmentation anormale du nombre de copies du gène dans la cellule, les copies surnuméraires se trouvant alors, soit sous forme intégrée dans un chromosome, soit sous forme de minichromosomes surnuméraires, les chromosomes double-minute (DM). Cette amplification entraîne généralement une augmentation du niveau de l'expression du gène.
- **Dérégulation de l'expression, stabilisation d'un m RNA**
Codant pour une oncoprotéine :
Les proto-oncogènes, lors de translocations chromosomiques, peuvent être déconnectés de leur environnement moléculaire normal et placés sous le contrôle inapproprié d'autres séquences à l'origine d'une modification de leur expression.

(Marlon Arnone and All, 2020)

1.3 Les Biothérapies

1.3. a Analyse génomique : In vivo (Cas cancer acquis) - In Silico de modélisation et prédiction

Analyse génomique : Des analyses de laboratoire cherchent maintenant à lire l'ADN des cellules malades permettant de définir les caractéristiques moléculaires d'une tumeur. Tel un code barre, cette analyse va permettre d'obtenir une signature génomique et ouvre la voie à la médecine personnalisée du cancer. L'analyse génomique permet de lire, en une seule manipulation de laboratoire, différentes zones de notre patrimoine génétique ou génome. Le génome correspond à

l'ensemble de notre matériel génétique. Il est situé dans le noyau de toutes les cellules de notre organisme et est constitué d'ADN.

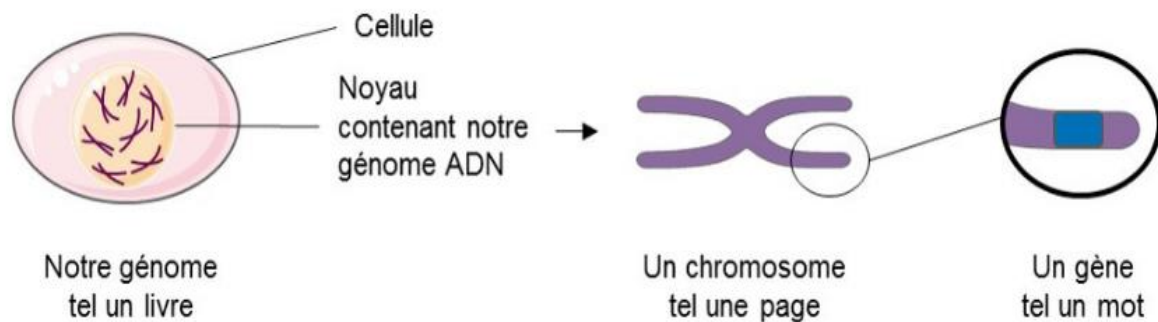


Figure 6 : Le génome

(Sylvie Gisselbrecht,2003)

Ainsi, tel un livre édité en un seul exemplaire, le génome regroupe toutes les informations qui caractérisent un individu. Tel des pages, il est composé de 23 paires de chromosomes, sur lesquels sont répartis près de 30 000 gènes. Et tels des mots, ces gènes expriment une à une les caractéristiques propres d'une personne (couleurs des yeux et de la peau, groupe sanguin, taille, hormones...) (Sylvie Gisselbrecht,2003)

1.3. b Tests génétiques – Séquençage – Profil génétique – Carte génétique

Pourquoi réaliser une analyse génomique ? L'analyse génomique a pour but de mieux caractériser la maladie permettant d'adapter et d'améliorer la prise en charge. Elle permet notamment de :

- prédire l'efficacité ou la résistance à un traitement, notamment par chimiothérapie et/ou thérapie ciblée;
- établir un diagnostic précis grâce à la caractérisation de la signature génomique ;
- prédire l'évolution de la maladie (agressivité de la tumeur, risque potentiel de métastases);
- permettre la prescription d'un traitement ciblé quand celui-ci est disponible.

(Marlon Arnone and All, 2020)

- **Une analyse génomique pour quel cancer ?** L'analyse génomique est en plein essor. Elle permet le développement de la médecine personnalisée du cancer. Le traitement est alors prescrit individuellement, en fonction de la signature génomique et des données biologiques de la tumeur ainsi que des caractéristiques du patient.
 - Différents tests sont réalisés pour éviter l'administration d'une chimiothérapie après chirurgie qui aurait été inutile pour certaines femmes atteintes d'un cancer du sein,

améliorant ainsi considérablement la qualité de vie tout en préservant les chances de guérison.

1.3. c Etudes in vivo épidémiologiques

L'analyse génomique en pratique est réalisée à partir d'un échantillon biologique (prise de sang, cellules tumorales...).

En laboratoire de biologie moléculaire, le matériel génétique contenu dans le noyau des cellules va être analysé par la technologie nommée « puce à ADN ». Celle-ci permet de visualiser le niveau d'activité d'un ensemble de gènes prédéfinis (biomarqueurs), reflétant ainsi la présence ou non d'altérations génétiques. Pour chaque gène, il sera possible de déterminer s'il est exprimé normalement, sous-exprimé, surexprimé ou encore éteint. La technologie de puce à ADN analyse simultanément l'expression de plusieurs dizaines de gènes.

- Une autre technique de séquençage à haut débit NGS (pour next generation sequencing) est également développée pour une analyse simultanée de plusieurs centaines de gènes. Ces deux techniques permettent d'obtenir les caractéristiques de la tumeur ou signature génomique.

1.3. d'Etudes in silico et biobanques : Détaillé dans la partie pratique expérimentale.

L'annotation des séquences génomiques consiste à extraire, à partir d'outils informatiques, le maximum d'information des données de séquences afin de prédire des caractéristiques phénotypiques permettant de guider le travail expérimental.

1.3. e Thérapie génique ciblée personnalisée La biotechnologie médicale, clé de l'innovation thérapeutique génique

Le développement des biotechnologies médicales et de la médecine personnalisée pourrait notamment jouer un rôle essentiel dans le traitement des cancers. Ici, des cellules cancéreuses.

Depuis le premier biomédicament commercialisé il y a moins de 30 ans, les biotechnologies se sont progressivement imposées comme un moteur majeur de la découverte de médicaments tout au long du cycle de leur développement et de leur production.

Tout d'abord, les biotechnologies ont permis de produire des médicaments que les méthodes industrielles classiques (extraction à partir d'organismes vivants, souvent des animaux, problèmes de purification, risque de contamination notamment virale) ne permettaient pas ou plus d'obtenir.

Par ailleurs, la disponibilité de protéines pures, en grande quantité, a permis le développement de très nombreux kits de diagnostic, alliant simplicité d'utilisation, sensibilité et spécificité, et a transformé l'approche de très nombreuses affections.

Ensuite, avec le séquençage du génome de nombreux organismes vivants et notamment celui de l'homme, l'identification de nouveaux gènes, l'étude de leurs polymorphismes facilitée par

l'utilisation de biopuces de plus en plus performantes, la découverte de nouveaux mécanismes moléculaires ont facilité la recherche de médicaments entièrement nouveaux.

Au-delà du médicament, de nouvelles perspectives thérapeutiques ; La médecine personnalisée se fait peu à peu une place dans le secteur des biotechnologies. Elle permet un meilleur ciblage de la maladie et un accompagnement plus efficace du patient grâce au recueil et à l'analyse des données.

La Thérapie génique des leucémies : Testée dès le début des années 1990 avec de nombreux essais tant chez l'animal que chez l'homme, la thérapie génique a donné naissance au concept de « l'ADN médicament ».

Des techniques dites de chirurgie du gène ou de saut d'exons, avec délétion de l'exon porteur d'une mutation et rétablissement du cadre de lecture du gène ont été testées avec succès chez l'animal (souris, chien) sur la dystrophine (maladies de Duchenne de Boulogne et de Becker) et sont en cours d'étude chez l'homme.

Les thérapies cellulaires bénéficient du développement très rapide des études sur les cellules souches qui peuvent se multiplier indéfiniment, donc à grande échelle et se différencier. Elles ouvrent la voie à une médecine régénératrice capable, par le remplacement de cellules détruites, de rétablir le fonctionnement d'un organe. (Sylvie Gisselbrecht, 2003)

1.3. f Les anti-oncogènes : gènes suppresseurs de tumeurs

Ces gènes sont aptes à inhiber la croissance cellulaire lorsqu'ils sont introduits par transfection dans les cellules tumorales. Cette propriété s'explique par la capacité de ces gènes à réguler négativement le cycle cellulaire et à induire l'apoptose ou mort cellulaire programmée. Action cellulaire récessive : une altération des 2 allèles est nécessaire à l'obtention d'une perte d'activité.

2 étapes sont nécessaires : 1ère étape somatique (cancer sporadique) ou germinale (cancer héréditaire : facteur de prédisposition).

Les altérations moléculaires à l'origine de la perte de fonction des gènes suppresseurs dans les tumeurs solides sont variées. Il peut s'agir de mutations ponctuelles, de délétions, d'insertions, d'anomalies de méthylation des promoteurs inhibant la transcription. **Mécanisme Oncogène 1.**

Biomarqueurs leucémiques pronostiques et prédictifs : Molécule biologique (enzyme, hormone...) dont la présence ou l'absence, ou encore la concentration dans les fluides biologiques ou les tissus est un indicateur d'un statut physiologique particulier.

Le choix des zones du génome à analyser sera fonction du type de cancer.

On parle alors de :

Biomarqueurs tumoraux pronostiques quand ils permettent de prédire l'évolution clinique de la maladie avant traitement. Ces biomarqueurs sont utilisés pour déterminer le risque potentiel de récurrence ou encore pour évaluer l'efficacité d'une chimiothérapie et/ou d'une thérapie ciblée.

Biomarqueurs tumoraux prédictifs quand ils permettent de prédire l'efficacité ou la résistance d'un traitement particulier. **(Seishi Ogawa, 2019)**

1.3. g Biomarqueurs leucémiques pronostiques et prédictifs

thérapeutiques futures, l'analyse génomique fera partie intégrante des examens inclus dans le parcours de soins du patient atteint de cancer. Ceci implique de connaître la signature génomique de chaque maladie et de pouvoir proposer une solution thérapeutique spécifique à chaque cas. Actuellement, le programme de recherche mondiale ICGC (pour International Cancer Genome Consortium) est en cours afin de caractériser « génomiquement » plus de 50 types de cancer.

En pratique, près de 25 000 génomes complets issus de cellules cancéreuses de patients sont séquencés et comparés. Les résultats devraient permettre d'identifier la signature génomique pour chaque maladie. Ces informations seront précieuses aussi bien pour un diagnostic rapide et précis, pour la prescription de thérapies ciblées, pour la prédiction de l'efficacité de certains traitements actuels que pour la participation à un essai clinique.

Ces analyses sont également le moyen d'identifier de nouvelles caractéristiques génomiques d'une maladie. Elles représentent de nouvelles cibles potentielles permettant de développer les médicaments de demain. Ceci explique pourquoi on parle de traitement personnalisé. **(Dohner H and All).**

CHAPITRE 2

Matériel & Méthodes

2.1 Matériel

- Matériel Non biologique

Ce présent travail est une étude in Silico réalisée pendant une période de 4 mois, du mois Mars 2021 à Juillet 2021.

Rappelons que notre objectif principale est l'identification in silico de quelques oncogènes grâce au modèle tumoral, cas des leucémies.

Pour répondre à la problématique de notre étude in Silico, nous avons utilisé des plateformes de bases de données qui utilisent la technologie de séquençage à haut débit comme les NGS (Next Generation Sequencing) à l'échelle internationale.

- Matériel biologique in silico

Ces données sont stockées dans des bases de données à accès libre (Open Source).

- Matériel non biologique :

Pour l'analyse de nos échantillons moléculaires nous avons utilisé des interfaces web, des logiciels, des softwares qui se base sur des algorithmes. Ces outils bioinformatiques sont répertoriés dans le tableau II.

2.2 Méthodes :

2.2.1 Principe, stratégie d'annotation et Protocol d'étude : Diagnostic de Prédiction préventif & Modélisation in silico des génomes

Notre approche de base simple et SMART d'immersion, de maîtrise et de validation de nos acquis de formation dans le sens des études in silico est fondée sur un protocole reparti en deux étapes de recherche, synthèse et d'étude :

- Recherche Bibliographique et synthèse documentaire
- Recherche et applications sur les biobanks : Interrogation des bases de données et plateformes bioinformatiques, alignement et modélisation des résultats pertinents de la synthèse bibliographique.
- Nos Etapes d'identification in silico des oncogènes leucémiques et leurs variants mutationnels

1. Recherche documentaire (Bibliographie d'études in vivo et épidémiologiques)

Listing des oncogènes leucémiques et des variants de chacun de ces oncogènes

2. Analyse & Modélisation BioInformatique :

L'annotation des séquences génomiques consiste à extraire, à partir d'outils informatiques, le maximum d'information des données de séquences afin de prédire des caractéristiques phénotypiques permettant de guider le travail expérimental.

- **Interrogation des bases de données & génération des données à haut débit**

Outils : NCBI – NIH- Blast Fasta

- a. Télécharger les données brutes à partir de NCBI et NIH : Interrogation des bases de données : La génération des données à haut débit et interrogation des bases de données
- b. Analyse des données
- c. Conclusion

- **En pratique : Démarche Globale**

- Etablir un protocole d'étude, une stratégie d'identification, et d'analyse in silico :
- Alignement des séquences et comparaison entre l'oncogène et ses variants
- Chercher la position de la mutation sur le codant du variant
- Recherche le codant de la séquence du gène oncogène et du variant sur NCBI
- Transformer la séquence en format fasta
- Alignement des séquences pour la comparaison sur NCBI blast en forma fasta
- Identification, prédiction de structure , de propriété, de fonction des variants des oncogènes

2.2.2 Méthodes de recherches documentaires :

Recherches d'articles et d'études en ligne.

2.2.3 Méthodes bioinformatiques :

- **Outil Orphanet**

Principe: Orphanet est un portail et serveur d'informations dédié aux maladies rares et aux médicaments orphelins en libre accès pour tous publics.

- **NCBI**
- **NIH**
- **Outil NCBI**

Principe : L'initiation de notre travail a sollicité l'Interrogation de la base de données NCBI (*National Cancer for Biology*). C'est une base de données d'Information de littérature biomédicale et de biomolécule. (<https://www.ncbi.nlm.nih.gov/>)

- **Outil Blast**

Principe : L'outil de recherche d'alignement local de base (BLAST) trouve les régions de similitude locale entre les séquences. Le programme compare les séquences de nucléotides ou de protéines aux bases de données de séquences et calcule la signification statistique des correspondances.

BLAST peut être utilisé pour déduire des relations fonctionnelles et évolutives entre les séquences ainsi que pour aider à identifier les membres des familles de gènes.

- **Blast Fasta**

Figure 7 : Protocol d'étude in silico

- Recherche de la séquence du gène oncogène et du variant sur Genbank
- Transformer la séquence en format fasta
- Alignement des séquences pour la comparaison sur NCBI blast en forma fasta

The figure displays several screenshots from the NCBI BLAST interface. The top-left screenshot shows the 'FASTA' input page for 'Homo sapiens chromosome 8, GRCh38.p13 Primary Assembly'. The top-right screenshot shows the 'BLAST' search results page with a table of sequences. The middle-left screenshot shows the 'Filter Results' section with options for organism, percent identity, and e-value. The middle-right screenshot shows the 'Sequences producing significant alignments' table. The bottom section shows a 'Graphic Summary' with a sequence alignment plot and a message: 'Putative conserved domains have been detected, click on the image below for detailed results.'

Descriptions **Graphic Summary** Alignments Taxonomy

hover to see the title click to show alignments Show Conserved Domains Alignment Scores < 40 40 - 50 50 - 80 80 - 200 >= 200 ?

100 sequences selected Putative conserved domains have been detected, click on the image below for detailed results.

RF #1 SuperFamilies

RF #3 SuperFamilies

Specific Hits SuperFamilies

Alignment view Taxonomy Download

100 sequences selected

Download GenBank Graphics

v-myc myelocytomatosis viral oncogene homolog, partial [Homo sapiens]
Sequence ID: CAL18312.1 Length: 287 Number of Matches: 1

Range 1: 15 to 267 GenBank Graphics

Score	Expect	Method	Identifiers	Positives	Gaps	Frame
382.04(982)	1e-117	Compositional matrix adjust.	252(252/99%)	252(252/99%)	0(252/0%)	+1
Query 2827			TRP LSG FTRMNDL DNVGAPVYCEIEEYVYDQVLEI QPAPSDCAKRE I L I Q	3806		
Subject 15			TRP LSG FTRMNDL DNVGAPVYCEIEEYVYDQVLEI QPAPSDCAKRE I L I Q	74		
Query 3807			ILGSGVGGGLGSPVYVTFSLRGRDGGGGSTADQLKRNTELLGGQWQDFICDP	3386		
Subject 75			ILGSGVGGGLGSPVYVTFSLRGRDGGGGSTADQLKRNTELLGGQWQDFICDP	134		
Query 3387			DDTF INKIIIDQGNVGSAAKLVSEKLVQARVQDSSGPRARHGVSTSLVYQ	3366		
Subject 135			DDTF INKIIIDQGNVGSAAKLVSEKLVQARVQDSSGPRARHGVSTSLVYQ	194		
Query 3367			DLGAAEEDIPVWVPPVPLDSSPSKSGDQNSAFSPSSQSLLSLSTESLQDQPPFLV	3546		
Subject 195			DLGAAEEDIPVWVPPVPLDSSPSKSGDQNSAFSPSSQSLLSLSTESLQDQPPFLV	254		
Query 3547			LHEETPTTSDS	3585		
Subject 255			LHEETPTTSDS	287		

Feedback

Download GenBank Graphics

truncated c-myc-P64 protein [Homo sapiens]
Sequence ID: AA48886.1 Length: 287 Number of Matches: 1

Range 1: 1 to 287 GenBank Graphics

Score	Expect	Method	Identifiers	Positives	Gaps	Frame
382.04(982)	7e-117	Compositional matrix adjust.	252(252/99%)	252(252/99%)	0(252/0%)	+1
Query 3878			TRP LSG FTRMNDL DNVGAPVYCEIEEYVYDQVLEI QPAPSDCAKRE I L I Q	3809		
Subject 1			TRP LSG FTRMNDL DNVGAPVYCEIEEYVYDQVLEI QPAPSDCAKRE I L I Q	68		
Query 3808			ILGSGVGGGLGSPVYVTFSLRGRDGGGGSTADQLKRNTELLGGQWQDFICDP	3389		
Subject 65			ILGSGVGGGLGSPVYVTFSLRGRDGGGGSTADQLKRNTELLGGQWQDFICDP	128		
Query 3388			DDTF INKIIIDQGNVGSAAKLVSEKLVQARVQDSSGPRARHGVSTSLVYQ	3369		
Subject 132			DDTF INKIIIDQGNVGSAAKLVSEKLVQARVQDSSGPRARHGVSTSLVYQ	188		
Query 3378			DLGAAEEDIPVWVPPVPLDSSPSKSGDQNSAFSPSSQSLLSLSTESLQDQPPFLV	3549		
Subject 181			DLGAAEEDIPVWVPPVPLDSSPSKSGDQNSAFSPSSQSLLSLSTESLQDQPPFLV	248		
Query 3508			LHEETPTTSDS	3588		
Subject 242			LHEETPTTSDS	253		

Feedback

Download GenBank Graphics

v-myc myelocytomatosis viral oncogene homolog, partial [Homo sapiens]
Sequence ID: CAL18312.1 Length: 287 Number of Matches: 1

Range 1: 15 to 267 GenBank Graphics

Score	Expect	Method	Identifiers	Positives	Gaps	Frame
382.04(982)	1e-117	Compositional matrix adjust.	252(252/99%)	252(252/99%)	0(252/0%)	+1
Query 2827			TRP LSG FTRMNDL DNVGAPVYCEIEEYVYDQVLEI QPAPSDCAKRE I L I Q	3806		
Subject 15			TRP LSG FTRMNDL DNVGAPVYCEIEEYVYDQVLEI QPAPSDCAKRE I L I Q	74		
Query 3807			ILGSGVGGGLGSPVYVTFSLRGRDGGGGSTADQLKRNTELLGGQWQDFICDP	3386		
Subject 75			ILGSGVGGGLGSPVYVTFSLRGRDGGGGSTADQLKRNTELLGGQWQDFICDP	134		
Query 3387			DDTF INKIIIDQGNVGSAAKLVSEKLVQARVQDSSGPRARHGVSTSLVYQ	3366		
Subject 135			DDTF INKIIIDQGNVGSAAKLVSEKLVQARVQDSSGPRARHGVSTSLVYQ	194		
Query 3367			DLGAAEEDIPVWVPPVPLDSSPSKSGDQNSAFSPSSQSLLSLSTESLQDQPPFLV	3546		
Subject 195			DLGAAEEDIPVWVPPVPLDSSPSKSGDQNSAFSPSSQSLLSLSTESLQDQPPFLV	254		
Query 3547			LHEETPTTSDS	3585		
Subject 255			LHEETPTTSDS	287		

Feedback

NCBI Protein

Full text: v-myc [v-myc]

Advanced Search Query Builder

Refinements

Crystal structure of the major quadruplex formed in the human c-myc promoter

3DHS: Full text: 18 4020556-4020556

Method: X-RAY DIFFRACTION 2.30 A

Organism: Homo sapiens

Macromolecule: DNA (5')

UniProt Link: 4

Figure 7 : Protocol d'étude in silico

- Recherche le codant de la séquence du gène oncogène et du variant sur Genbank
- Transformer la séquence en format fasta
- Alignement des séquences pour la comparaison sur NCBI blast en forma fasta

CHAPITRE 3

Résultats & Discussions

3.1 Résultats :

- Résultats Bibliographiques
 - Résultats bioinformatiques
- Cas pratiques d'identification d'oncogènes et variants leucémiques in silico : Recherche documentaire et in silico.**

Les Principaux réarrangements chromosomiques décrits dans les leucémies humaines. Les réarrangements peuvent conduire à:

- Soit à la surexpression de protéines oncogéniques (A)
- Soit à la production de protéines de fusion (B)

Dans notre mémoire nous allons étudier les variants des oncogènes de la catégorie (A) relatifs à la surexpression de protéines oncogéniques, car les gènes oncogènes de la catégorie B ne possèdent pas de variants.

Les 05 Oncogènes objets de notre étude sont donc les suivants :

- 1. MYC (8q34) t(8;14)(q24;q11) : Variants (t(8;14)(q24;q32) , t(2;8)(p12;q24), t(8;22)(q24;q11))
- 2. HOX11(10q24) t(10;14)(q24;q11) : Variant t(7;10)(q35;q24)
- 3. LMO2(11p13) t(11;14)(p13;q11) : Variant t(7;11)(q35;p13)
- 4. EVI 1(3;q26) t(3;3)(q21;q26) : Variant inv (3)(q21;q26)
- 5. LAZ 3/BCL6 (3q27) t(3;14)(q27;q32) : Variants (t(3;22)(q27;q11), t(3;4)(q27;p11))

a. Listing des oncogènes leucémiques et des variants de chacun de ces oncogènes

Tableau. Principaux réarrangements chromosomiques décrits dans les leucémies humaines. Les réarrangements peuvent conduire à:

- Soit à la surexpression de protéines oncogéniques (A)
- Soit à la production de protéines de fusion (B)

A. SUREXPRESSION DE PROTÉINES ONCOGÉNIQUES										
Type de gène	Fonction	Gène	Translocation	Régions régulatrices	Type de leucémie	Gènes Variants	Translocation	Région régulatrice Variants	Type de leucémie	
Facteurs de transcription et régulateurs transcriptionnels	Protéine de type bHLH	MYC (8q34)	t(8;14)(q24;q11)	TCR α	LAL T	03 Variants MYC (8q34)	t(8;14)(q24;q32) t(2;8)(p12;q24) t(8;22)(q24;q11)	IgH, IgL	Lymphomes et leucémies de Burkitt, leucémies lymphoïdes chroniques	
		TAL 1 (1p32)	t(1;14)(p32;q11)	TCR α	LAL T	Pas de variants	/			
		LYL 1 (19p13)	t(7;19)(q35;p13)	TCR β	LAL T					
	Protéine à homéodomaine	HOX11 (10q24)	t(10;14)(q24;q11)	TCR α, TCR β			01 Variant de HOX11 (10q24)	t(7;10)(q35;q24)	TCR α, TCR β	LAL T
		HOX11L2 (5q35)	t(5;14)(q35;q32)	gène: CTIP2	LAL T	Pas de variants	/			
	Protéine à domaine LIM	LMO1 (11p15)	t(11;14)(p15;q11)	TCR δ	LAL T	Pas de variants	/			
LMO2 (11p13)		t(11;14)(p13;q11)	TCR δ, TCR α, TCR β	LAL T	01 Variant de LMO2 (11p13)	t(7;11)(q35;p13)	TCR δ, TCR α, TCR β	LAL T		
Protéine à doigts de zinc	EVI 1 (3; q26)	t(3;3)(q21;q26)	gène: ribophorine	LAM	01 Variant de EVI 1 (3; q26)	t(3;3)(q21;q26)	gène: ribophorine	LAM		
	LAZ 3/BCL6 (3q27)	t(3;14)(q27;q32)	IgH, IgL	Lymphomes non hodgkiniens	02 Variants LAZ 3/BCL6 (3q27)	t(3;22)(q27;q11) t(3;4)(q27;p11)	IgH, IgL	Lymphomes non hodgkiniens		

1. MYC (8q34) t(8;14)(q24;q11) : Gène de protéine de type bHLH, facteur de transcription et régulateur transcritionnel

: Gène de protéine de type bHLH, facteur de transcription et régulateur transcritionnel

The screenshot shows the NCBI website interface. At the top, there is a search bar for nucleotide sequences. Below it, a COVID-19 information banner is visible. The main content area displays the 'Homo sapiens chromosome 8, GRCh38.p13 Primary Assembly' with the NCBI Reference Sequence: NC_000008.11. A 'FASTA' tab is selected, showing the full nucleotide sequence. On the right side, there are options to 'Change region shown' (set to 'Selected region' from 127735434 to 127742951) and 'Customize view'. There are also links to 'Analyze this sequence', 'Run BLAST', 'Pick Primers', 'Highlight Sequence Features', 'Find in this Sequence', and 'Show in Genome Data Viewer'.

The screenshot shows the NCBI BLAST search interface. The browser address bar indicates the URL: blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome. The page header includes the NIH logo and 'U.S. National Library of Medicine National Center for Biotechnology Information'. A COVID-19 information banner is present. The main heading is 'BLAST® >> blastx' with a sub-heading 'Translated BLAST: blastx'. The 'Enter Query Sequence' section contains a text area with the following sequence: GCTATTGCTGTCTAATTA, CCTCATTGTCTCAGTCTCAAAAGTAGGCTTCAGCTCCCTGACTTTGGGAT, TTTAATCTACCAACCCA, TAAATCAATAAATAATTCTTTCTTGA. Below the text area are fields for 'Query subrange' (From and To), 'Or, upload file', 'Genetic code' (Standard (1)), and 'Job Title' (NC_000008.11:127735434-127742951 Homo sapiens...). There is also a checkbox for 'Align two or more sequences'. The 'Choose Search Set' section shows the 'Database' set to 'Non-redundant protein sequences (nr)'. A 'New columns added to the Description Table' notification is visible on the right.

BLAST® » blastx » results for RID-DVHUE5XP013 Home Recent Results Saved Strategies Help

[Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title **NC_000008.11:127735434-127742951 Homo sapiens...**
 RID **DVHUE5XP013** Search expires on 07-02 21:46 pm [Download All](#) ▼
 Program **BLASTX** [Citation](#) ▼
 Database **nr** [See details](#) ▼
 Query ID **lc|Query_64541**
 Description **NC_000008.11:127735434-127742951 Homo sapiens chr ...**
 Molecule type **dna**
 Query Length **7518**
 Other reports [?](#)

Filter Results

Organism only top 20 will appear exclude
 Type common name, binomial, taxid or group name
[+ Add organism](#)

Percent Identity to E value to Query Coverage to

[Filter](#) [Reset](#)

Descriptions Graphic Summary Alignments Taxonomy [Feedback](#)

Other reports [?](#) [Filter](#) [Reset](#)

Descriptions Graphic Summary Alignments Taxonomy

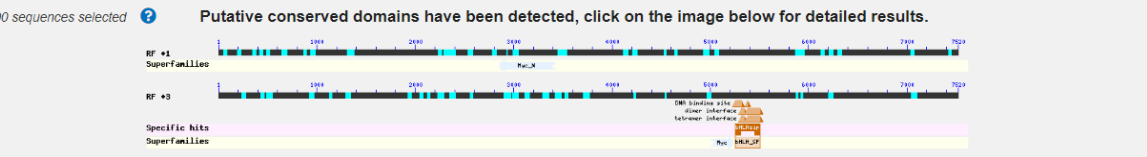
Sequences producing significant alignments Download ▼ **New** Select columns ▼ Show [?](#)

select all 100 sequences selected [GenPept](#) [Graphics](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> v-myc myelocytomatosis viral oncogene homolog [Homo sapiens]	Homo sapiens	383	383	10%	6e-117	99.21%	267	CAL18326.1
<input checked="" type="checkbox"/> truncated c-myc-P64 protein [Homo sapiens]	Homo sapiens	382	382	10%	7e-117	100.00%	257	AAA88095.1
<input checked="" type="checkbox"/> v-myc myelocytomatosis viral oncogene homolog [Homo sapiens]	Homo sapiens	382	382	10%	8e-117	99.60%	267	CAL18312.1

Descriptions **Graphic Summary** Alignments Taxonomy

[hover to see the title](#) [click to show alignments](#) Show Conserved Domains Alignment Scores < 40 40 - 50 50 - 80 80 - 200 >= 200 [?](#)



Alignment view [Restore defaults](#) Download ▼

100 sequences selected [?](#) [Next](#) [Previous](#) [Descriptions](#)

[Download](#) ▼ [GenPept](#) [Graphics](#)

v-myc myelocytomatosis viral oncogene homolog, partial [Homo sapiens]
 Sequence ID: [CAL18326.1](#) Length: 267 Number of Matches: 1

Range 1: 15 to 267 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
383 bits(983)	6e-117	Compositional matrix adjust.	251/253(99%)	251/253(99%)	0/253(0%)	+1
Query 2827	TMPLNVSFTNRRNYDLDYDVSQPYFYCDEEENFYqggggSELQPPAPSEDIIKKFELLpTP					3006
Sbjct 15	TMPLNVSFTNRRNYDLDYDVSQPYFYCDEEENFYQ0000SELQPPAPSEDIIKKFELL TP					74
Query 3007	pLSPRRSGLCSPSYVAVTPFSLRGDNDGGGGSFSTADQLEHMTTELLGGDMVNSFICDP					3186
Sbjct 75	SLSPRRSGLCSPSYVAVTPFSLRGDNDGGGGSFSTADQLEHMTTELLGGDMVNSFICDP					134
Query 3187	DDETFIKNIIIQDCMWSGFSAAAKLVSEKLASVQAARKDGSFPNARGHSVCSTSSLYLQ					3366
Sbjct 135	DDETFIKNIIIQDCMWSGFSAAAKLVSEKLASVQAARKDGSFPNARGHSVCSTSSLYLQ					194
Query 3367	DLAAAASECIDPSVVFYPYPLNDSSSPKscasqdsafspssdLLsstesspQGSPEPLV					3546
Sbjct 195	DLAAAASECIDPSVVFYPYPLNDSSSPKSCASQDSSAFSPSSDLSLSTESSPQGSPEPLV					254
Query 3547	LHEETPPTTSSDS 3585					
Sbjct 255	LHEETPPTTSSDS 267					

Related Information
[Gene](#) - associated gene details



[Feedback](#)

Download ▾ GenPept Graphics ▾ Next ▲ Previous ◀ Descriptions

truncated c-myc-P64 protein [Homo sapiens]
 Sequence ID: [AAA88095.1](#) Length: 257 Number of Matches: 1
[See 2 more title\(s\)](#) ▾ [See all Identical Proteins \(IPG\)](#)

Range 1: 1 to 253 [GenPept](#) [Graphics](#) ▾ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
382 bits(981)	7e-117	Compositional matrix adjust.	253/253(100%)	253/253(100%)	0/253(0%)	+1
Query 2830		MPLNVSFTNRNYLDYDSVQPYFYCDEEENFYqqqqqSELOPPAPSEDIWKKFELLptpp				3009
Sbjct 1		MPLNVSFTNRNYLDYDSVQPYFYCDEEENFYQ0000SELOPPAPSEDIWKKFELLTPP				60
Query 3010		lpsrnrsglcspsYVAVTPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNSQSFICDPD				3189
Sbjct 61		LSPSRRLGCLCSPSYVAVTPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNSQSFICDPD				120
Query 3190		DETFIKNIIIQDCMWSGFSAAAKLVSEKLASVQAARKDGSFPNARGHVCSTSSLYLQD				3369
Sbjct 121		DETFIKNIIIQDCMWSGFSAAAKLVSEKLASVQAARKDGSFPNARGHVCSTSSLYLQD				180
Query 3370		LSAAASECIDPSVVFYPPLNDSSSPKscasqdsafspssdsllsstesspqspeplVL				3549
Sbjct 181		LSAAASECIDPSVVFYPPLNDSSSPKSCASQDSSAFSPSSDLSLSTESSPQGSPEPLV				240
Query 3550		HEETPPTTSSDSG 3588				
Sbjct 241		HEETPPTTSSDSG 253				

Feedback

Download ▾ GenPept Graphics ▾ Next ▲ Previous ◀ Descriptions

v-myc myelocytomatosis viral oncogene homolog, partial [Homo sapiens]
 Sequence ID: [CAL18312.1](#) Length: 267 Number of Matches: 1

Range 1: 15 to 267 [GenPept](#) [Graphics](#) ▾ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
382 bits(982)	8e-117	Compositional matrix adjust.	252/253(99%)	252/253(99%)	0/253(0%)	+1
Query 2827		TMPLNVSFTNRNYLDYDSVQPYFYCDEEENFYqqqqqSELOPPAPSEDIWKKFELLtp				3006
Sbjct 15		TMPLNVSFTNRNYLDYDSVQPYFYCDEEENFYQ0000SELOPPAPSEDIWKKFELL TP				74
Query 3007		p1psrnrsglcspsYVAVTPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNSQSFICDP				3186
Sbjct 75		PLSPSRRLGCLCSPSYVAVTPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNSQSFICDP				134
Query 3187		DDETFIKNIIIQDCMWSGFSAAAKLVSEKLASVQAARKDGSFPNARGHVCSTSSLYLQ				3366
Sbjct 135		DDETFIKNIIIQDCMWSGFSAAAKLVSEKLASVQAARKDGSFPNARGHVCSTSSLYLQ				194
Query 3367		DLAAASECIDPSVVFYPPLNDSSSPKscasqdsafspssdsllsstesspqspeplVL				3546
Sbjct 195		DLAAASECIDPSVVFYPPLNDSSSPKSCASQDSSAFSPSSDLSLSTESSPQGSPEPLV				254
Query 3547		LHEETPPTTSSDS 3585				
Sbjct 255		LHEETPPTTSSDS 267				

Feedback

Figure 8 : Protocol d'étude in silico de l'oncogène MYC

1.1. Variants du MYC (8q34) t(8;14)(q24;q11):

t(8;14)(q24;q32).

t(2;8)(p12;q24).

t(8;22)(q24;q11).

- Protéine du MYC (8q34) t(8;14)(q24;q11):

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation MyPDB

PDBe-101 PDBe EMDataResource Worldwide Protein Data Bank

Celebrating 50 YEARS OF Protein Data Bank

Search History Browse Annotations MyPDB

QUERY: Full Text = "MYC (8q34)" Open In Query Builder JSON MyPDB Login

Advanced Search Query Builder Help

Refinements Clear All

Summary Gallery Compact -- Tabular Report -- Score Download Files All Selected

Displaying 1 to 25 of 1049 Structures Page 1 of 42 Previous Next Display 25 per page

SCIENTIFIC NAME OF SOURCE ORGANISM Clear

- Homo sapiens (910)
- Mus musculus (74)
- Danio rerio (37)
- synthetic construct (27)
- Rattus norvegicus (21)
- Escherichia coli K-12 (10)
- Arabidopsis thaliana (9)
- Bos taurus (6)
- Saccharomyces cerevisiae (6)
- Thermus thermophilus HB8 (6)

More...

TAXONOMY Clear

3D View

6AU4 Download File View File

Crystal structure of the major quadruplex formed in the human c-MYC promoter

Stump, S., Mou, T.C., Sprang, S.R., Natale, N.R., Beall, H.D.

(2018) PLoS One 13: e0205584-e0205584

Released 2018-09-12

Method X-RAY DIFFRACTION 2.35 Å

Organisms Homo sapiens

Macromolecule DNA (5'-D(*TP*GP*AP*GP*GP*TP*GP*GP*GP*TP*AP*GP*GP*TP*GP*GP*TP*AP*A)-3') (nucleic acid)

Unique Ligands K

2. HOX11(10q24) t(10;14)(q24;q11)

Gene:

COVID-19 Information
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

Full Report

TLX1 T cell leukemia homeobox 1 [*Homo sapiens* (human)]

Gene ID: 3195, updated on 11-Jun-2021

Summary

Official Symbol TLX1 provided by HGNC
Official Full Name T cell leukemia homeobox 1 provided by HGNC
Primary source [HGNC:HGNC:5056](#)
See related [Ensembl:ENSG00000107807](#) [MIM:186770](#)
Gene type protein coding
RefSeq status REVIEWED
Organism [Homo sapiens](#)
Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as TCL3; HOX11
Summary This gene encodes a nuclear transcription factor that belongs to the NK-linked or NK-like (NKL) subfamily of homeobox genes. The encoded protein is required for normal development of the spleen during embryogenesis. This protein is also involved in specification of neuronal cell fates. Ectopic expression of this gene due to chromosomal translocations is associated with certain T-cell acute

Table of contents

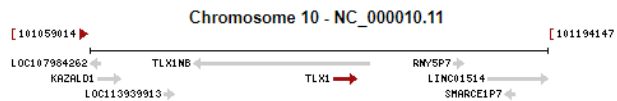
- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Expression
- Bibliography
- Phenotypes
- Variation
- Interactions
- General gene information
 - Markers, Clone Names, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links

Genomic context

Location: 10q24.31

Exon count: 4

Annotation release	Status	Assembly	Chr	Location
109.20210514	current	GRCh38.p13 (GCF_000001405.39)	10	NC_000010.11 (101130773..101137789)
105.20201022	previous assembly	GRCh37.p13 (GCF_000001405.25)	10	NC_000010.10 (102891057..102897546)



NCBI Resources

Nucleotide

COVID-19 Information
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

FASTA

Homo sapiens chromosome 10, GRCh38.p13 Primary Assembly
 NCBI Reference Sequence: NC_000010.11
[GenBank](#) [Graphics](#)

>NC_000010.11:101130773-101137789 Homo sapiens chromosome 10, GRCh38.p13 Primary Assembly
 CTGGGCTCCATACCTGGGGAGTAGACAGTCTCTACTTTCTAAAAGCTGGACTTGAATTTGGAGTAAA
 TCTCCGGTTGAGTACAGGTTGTTTCGACGCTGAGCCCTTGAGGAGATCTCCAGTTGGGACAGAGATC
 CCTTCTCAGACGCTTTGTTGGCTGGACTCCTTTGGCCAGTTCAAAGTGAAGGAGGGCTCCCAACAGG
 CCGGGAAGACAGTTGACTTCACCCCTCCTTGGGTTTGTCTGTCTGTCCGCTCTCGGAAATGGTGGCTTCT
 GTTTTCCCTTTTCTTTAAGCCTCGCCTTGTTCCTCTCTCTCTCTCATGAACACTCCGAGTCTT
 GGCTCCGCTCCCTATCTCTGGCTCTGATCTGTCTGGCTTCTGGCTTCTCTCTCCCTCCCTCC
 CCTCCTCGCTGTCATTCACCCGCTCCTCTCCGCGCACAGCCATGGAGAGACCAGTCSAAACCCG
 GAAGCTCTTTCGACCGGCTTTTTCGCTGGTGTGATGTCCAGAGTCAACAGGAGGAGGAGCCG
 GAGCGGGAAGCAGAAAGCAGAGAGGGAAGATACGGCCCTCTCTCCCTCCCTCCCTCCCTCTAC
 TTTAGCTTTCTCGCACTTCGCTTCAAAGTCTCCGCGCAGCAGBAGCCGCTGTTGCTCCAGCCCT
 GCTAGCTGCCCCGAGCGCAGCGCAGCGCAGCCGCGCCGCGCCCGCCGCTGGGCGCAGGCGCAGCA
 TGGAGACCTGGGTCGACACCTCCACCCGGTCAAGCAGAGCCATTAGCTTCGGCATCGACAGAT
 CCTCAACAGCCGAGCAGGTTGGCTGATGGGACCCGCTCGCCCTCCAGGACGGGAATACGGCTT
 GGCTGCTTGGTCGAGGCGCTACACTTACGGCGCGGGGGCTCCGCGCGCGCAGCGGGGCTGAGGAG
 CCGGGGCTATGGTACTGAGAGCCCGCGCGCCCGGAGGCGCCGCGAGCGCGCGCGCTGACAGAT
 GGGTCTTGCAGCGCTCTACAAGTGAACATGGCTTGGCAGCGCGCCCGTCTGGCGCGCGCGG
 GCGCAGCGCGCGGCTCGGGGCACTACGCGCTCGCGGGGTAATCCGGTGCAGCAGCGCGCTCG

Change region shown

Whole sequence
 Selected region
 from: 101130773 to: 101137789

Customize view

Analyze this sequence

- Run BLAST
- Pick Primers
- Highlight Sequence Features
- Find in this Sequence
- Show in Genome Data Viewer

COVID-19 Information

Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

BLAST® » blastx » results for RID-DVK8E9GY013

Home Recent Results Saved Strategies Help

[< Edit Search](#) [Save Search](#) [Search Summary](#)

[How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title **NC_000010.11:101130773-101137789 Homo sapiens...**
 RID [DVK8E9GY013](#) Search expires on 07-02 22:11 pm [Download All](#)
 Program BLASTX [Citation](#)
 Database nr [See details](#)
 Query ID lc|Query_93806
 Description NC_000010.11:101130773-101137789 Homo sapiens chr...
 Molecule type dna
 Query Length 7017
 Other reports

Filter Results

exclude

Organism only top 20 will appear

Type common name, binomial, taxid or group name

[+ Add organism](#)

Percent Identity to E value to Query Coverage to

[Filter](#) [Reset](#)

Descriptions Graphic Summary Alignments Taxonomy

[Feedback](#)

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments Download [New](#) Select columns Show 100

select all 100 sequences selected [GenPept](#) [Graphics](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> T-cell leukemia homeobox protein 1 [Papio anubis]	Papio anubis	232	451	17%	1e-62	87.04%	412	XP_003904190.1
<input checked="" type="checkbox"/> T-cell leukemia homeobox protein 1 isoform X1 [Theropithecus gelada]	Theropithecus gelada	228	324	14%	3e-62	85.56%	339	XP_025252984.1
<input checked="" type="checkbox"/> PREDICTED: T-cell leukemia homeobox protein 1 isoform X1 [Mandrillus leucophaeus]	Mandrillus leucophaeus	228	325	14%	3e-62	86.08%	342	XP_011832811.1

Other reports [Filter](#) [Reset](#)

Descriptions Graphic Summary **Alignments** Taxonomy

Alignment view Pairwise [Restore defaults](#) Download

100 sequences selected

[Download](#) [GenPept](#) [Graphics](#) Sort by: E value [Next](#) [Previous](#) [Descriptions](#)

T-cell leukemia homeobox protein 1 [Papio anubis]
 Sequence ID: [XP_003904190.1](#) Length: 412 Number of Matches: 3

Range 1: 4 to 273 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
232 bits(592)	1e-62	Compositional matrix adjust.	235/270(87%)	243/270(90%)	3/270(1%)	+2
Query 542		NSERAAGAGKQKPERGRIRRp1spspstTLAFRLTSLPSLRAARSRCCLpapascpps				721
Sbjct 4		+ +RA+ + + + G+ + PLSPSPPTLA LRTSLPSLRAARSRCCLPAPASCP SR				63
Query 722		aqrapppppppgarasMEhlgphh1hpgHAEPISFGIDQILNSPDQGGCMGPASRLQDGE				901
Sbjct 64		AQRAPPPGPP GA ASMEHLGPHLHPGHAEPISFGIDQILNSPDQGGCMGPASRLQDGE				123
Query 902		YLGCLVggaytygggsaaatgaggagaygtggggppagggga---CSMGPLTGSY				1072
Sbjct 124		YLGCLVGGAYTYGGGSAPGAGAGGAGAYGTGGPGGGPAGGGGACSMGLAGSY				183
Query 1073		NVMalagppggggssggagalsaaGVIRVPAHRPLAGAVHPQPLATGLPTVPSVP				1252
Sbjct 184		NVMALAGGPGGGG SGGAGALSAAGVIRVPAHRPL GAVAHPOPLATGLPTVPSVP				243

Related Information
[Gene](#) - associated gene details
[Genome Data Viewer](#) - aligned genomic context

[Feedback](#)

Download GenPept Graphics Sort by: E value Next Previous Descriptions

T-cell leukemia homeobox protein 1 isoform X1 [Theropithecus gelada]
 Sequence ID: [XP_025252984.1](#) Length: 339 Number of Matches: 2

Range 1: 4 to 273 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
228 bits(582)	3e-62	Compositional matrix adjust.	231/270(86%)	239/270(88%)	3/270(1%)	+2
Query 542	NSERAAGAGKQKPERGRIRRp1spspssTLAFLRTSLPSLRAARSRCCLpapascppsr					721
Sbjct 4	+ +RA+ + + + G+ + PLSPPSSTLA LRTSLPSLRAAR RCCLPAPASC SR					63
Query 722	aqrapppppppggarasMEhlgphh1hphgAEPISFGIDQILNSPDQGGCMGPASRLQDGE					901
Sbjct 64	AQ APPPPPP GA ASMEHLGPHHLHPGHAEPISFGIDQILNSPDQGGCMGPASRLQDGE					123
Query 902	YLGCLVggaytyggggsaaatgaggagaytgppppppagggga--CSMGPLTGSY					1072
Sbjct 124	YLGCLVGGAYTYGGGSA GA GAGAYTGGPPG G G G CSMGPL GSY					183
Query 1073	NVMalagppppppggggsaggagalsaaagVIRVPAHRPLAGAVAHPOPLATGLPTVPSV					1252
Sbjct 184	NVMALAGGPPGGGG SGGAGALSAAGVIRVPAHRPL GAVAHPOPLATGLPTVPSV					243
Query 1253	AMPGVNLLTGLTFPMESNRRYTKDRFTGE 1342					
Sbjct 244	AMPGVNLLTGLTFPMESNRRYTKDRFTGH 273					

Feedback

Download GenPept Graphics Sort by: E value Next Previous Descriptions

PREDICTED: T-cell leukemia homeobox protein 1 isoform X1 [Mandrillus leucophaeus]
 Sequence ID: [XP_011832811.1](#) Length: 342 Number of Matches: 2

Range 1: 4 to 276 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
228 bits(582)	3e-62	Compositional matrix adjust.	235/273(86%)	243/273(89%)	6/273(2%)	+2
Query 542	NSERAAGAGKQKPERGRIRRp1spspssTLAFLRTSLPSLRAARSRCCLpapascppsr					721
Sbjct 4	+ +RA+ + + + G+ + PLS SPPSSTLA LRTSLPSLRAARSRCCLPAPASC SR					63
Query 722	aqrapppppppggarasMEhlgphh1hphgAEPISFGIDQILNSPDQGGCMGPASRLQDGE					901
Sbjct 64	AQ APPPPPP GA ASMEHLGPHHLHPGHAEPISFGIDQILNSPDQGGCMGPASRLQDGE					123
Query 902	YLGCLVggaytyggggsaaatgaggagaytgppppppagggga----CSMGPLT					1063
Sbjct 124	YLGCLVGGAYTYGGGSA GAGGAGAYTGGPPGGP G GG CSMGPL					183
Query 1064	GSYVNMAlagppppppggggsaggagalsaaagVIRVPAHRPLAGAVAHPOPLATGLPTVP					1243
Sbjct 184	GSYVNMALAGGPPGGGG SGGAGALSAAGVIRVPAHRPL GAVAHPOPLATGLPTVP					243
Query 1244	SVPAMPGVNLLTGLTFPMESNRRYTKDRFTGE 1342					
Sbjct 244	SVPAMPGVNLLTGLTFPMESNRRYTKDRFTGH 276					

Feedback

Figure 9 : Protocol d'étude in silico de l'oncogène HOX11

2.1. Variants du HOX11(10q24) t(10;14)(q24;q11)

- Variant t(7;10)(q35;q24) du HOX11(10q24) t(10;14)(q24;q11)
- Protéine du HOX11(10q24) t(10;14)(q24;q11)

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation MyPDB

PDB-101 PDB PROTEIN DATA BANK EMDB RESOURCES Protein Data Bank for EMDB Nucleic Acid Database Protein Data Bank Foundation

Search History Browse Annotations MyPDB

QUERY: Full Text = "MYC (8q34)" Open In Query Builder JSON MyPDB Login

Advanced Search Query Builder Help

Refinements Clear All

Summary Gallery Compact -- Tabular Report -- Score Download Files All Selected

Displaying 1 to 25 of 1049 Structures Page 1 of 42 Previous Next Display 25 per page

6AU4 Download File View File

Crystal structure of the major quadruplex formed in the human c-MYC promoter
Stump, S., Mou, T.C., Sprang, S.R., Natale, N.R., Beall, H.D.
(2018) PLoS One 13: e0205584-e0205584

Released 2018-09-12
Method X-RAY DIFFRACTION 2.35 Å
Organisms Homo sapiens
Macromolecule DNA (5'-D(TP*GP*AP*GP*GP*GP*TP*GP*GP*GP*TP*AP*GP*GP*GP*TP*GP*GP*TP*AP*A)-3') (nucleic acid)
Unique Ligands K

SCIENTIFIC NAME OF SOURCE ORGANISM Clear

- Homo sapiens (910)
- Mus musculus (74)
- Danio rerio (37)
- synthetic construct (27)
- Rattus norvegicus (21)
- Escherichia coli K-12 (10)
- Arabidopsis thaliana (9)
- Bos taurus (6)
- Saccharomyces cerevisiae (6)
- Thermus thermophilus HB8 (6)

More...

TAXONOMY Clear

- Eukaryota (1015)

3D View

Contact Us

3. LMO2(11p13) t(11;14)(p13;q11):

NCBI Resources How To Sign in to NCBI

Gene LMO2(11p13) Search

COVID-19 Information

Gene sources: Genomic, Categories, Sequence content, Status: Current

Search results: 2 items

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> LMO2 ID: 4005	LIM domain only 2 [<i>Homo sapiens</i> (human)]	Chromosome 11, NC_000011.10 (33858576..33892076, complement)	LMO-2, RBTN2, RBTN1, RHOM2, TTG2	180385
<input type="checkbox"/> RCN1 ID: 5954	reticulocalbin 1 [<i>Homo sapiens</i> (human)]	Chromosome 11, NC_000011.10 (32091074..32105722)	HEL-S-84, PIG20, RCAL, RCN	602735

LMO2 LIM domain only 2 [*Homo sapiens* (human)]

Gene ID: 4005, updated on 11-Jun-2021

Summary

Official Symbol LMO2 provided by HGNC

Official Full Name LIM domain only 2 provided by HGNC

Primary source HGNC:HGNC:6642

See related Ensembl:ENSG00000135363 MIM:180385

Gene type protein coding

RefSeq status REVIEWED

Organism [Homo sapiens](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as TTG2; LMO-2; RBTN2; RHOM2; RBTN1

Summary LMO2 encodes a cysteine-rich, two LIM-domain protein that is required for yolk sac erythropoiesis. The LMO2 protein has a central and crucial role in hematopoietic development and is highly conserved. The LMO2 transcription start site is located approximately 25 kb downstream from the 11p13 T-cell translocation cluster (11p13 ttc), where a number T-cell acute lymphoblastic leukemia-specific translocations occur. Alternative splicing results in multiple transcript variants encoding different isoforms.[provided by RefSeq, Nov 2008]

Table of contents: Summary, Genomic context, Genomic regions, transcripts, and products, Expression, Bibliography, Phenotypes, Variation, Interactions, General gene information, General protein information, NCBI Reference Sequences (RefSeq), Related sequences, Additional links

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Help

COVID-19 Information
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

FASTA Send to: **Change region shown**
 Whole sequence
 Selected region
 from: 33858576 to: 33892076 **Update View**

Homo sapiens chromosome 11, GRCh38.p13 Primary Assembly
 NCBI Reference Sequence: NC_000011.10
[GenBank](#) [Graphics](#)

>NC_000011.10:c33892076-33858576 Homo sapiens chromosome 11, GRCh38.p13 Primary Assembly
 AAACCTGGATTGCAACATCTCAAGAACAGCAACGGGCTTGAGCAGAAATGAGAAGGAAATACCCC
 CACCTGCCCTCAGCCGTTAAGTGGGCTTTGCTATTCAAAAGGGCTCTGGGTGCTCGGAGAGAGGGGA
 GATGGCACAGGCACCAAGGTGCTAGGGTGCAGGGCTCCCGAAGAGGAAACAGGTGCAAAAGCAGCAATTAG
 GCCCAAGAGGTATCCGTGGGCAAGCAGCTAGATCGATGGGGAAGCCACAGGATTAATCATCTGCTG
 TGGTAGTAGGCTTCAATTAATCTCTGATGAATGGACGATTGCAAGGAACTTTTTCACTCTCAAGGAG
 CCAGAAGAGTGGTGAATTAATGGTCTTTAAAAAAGAAAGACTCCAAGGGGTACAAGCTTCAAGCT
 TCCCTGTGGTGTTCGGGCACTCTGGCTCTTTGGAGGGGGAAGAGCATCTAATCATAGGCTGTG
 ACTTGGAGAGGCCACAGGTGAGTCAAGGCGCTTGTGTGTGAGCCGCTGGTTGCTAATGTCTCGGGAAC
 TGCTAATTACCATCTAATTTGTTCTGTACACTGGCCACAAGGGAGACCACTCAATTATCTGATCC
 AAAAGAGAGGCTGTGTTTACCATTAATCTCCCAAGCACTGTGTGAGTGTGTGCAATGTGTGTGTGTG
 GT
 CTGCCCTTGTAGCATTTCACACACATCCAGTTCTGAGTTGAGGGCAAGATTGTAATGGGGTCTTC
 AGAAGACTTGGCTCCCAAGGAGGTTAATCACTCAAGGGGTGTTTTGAAAAACAAGGGGATGAGGA
 AAGTGGCTACAAGTTCCCTCCATCACAGAGGATCAATGAAGGGGAACAAGATTGGAGCATCTACTTT
 GTGCCAGGACTGTGCTGGGTTTACACGCTCTCTTGTGACTTGGATAACTCCCTTGAAGCAG
 ATATTACTCAGGTTCTTGGCGTAAATGTTTCAGCTTAAGAAACCAACCACTACTTTTAGGTTGCC
 TGAAGAGTGAATCTCTGTGAGGCTTTTCTGTCTTTATCTGTTAAATAAATTAAGGCTGTGTTGTG
 CTCCGAAATACACTTATAGCACTTAAAGATGAGCAGCTCAGTTTCTCAITTTGTAATATAGGGGTAC

Customize view
Display options
 Show reverse complement
 Show gap features **Update View**

Analyze this sequence
 Run BLAST
 Pick Primers
 Highlight Sequence Features

BLAST » blastx » results for RID-DVZE07CH013 Home Recent Results Saved Strategies Help

[Edit Search](#) Save Search Search Summary [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

There was a problem with the search. Please, contact [Help Desk](#) and include RID DVZE07CH013.
 CPU usage limit was exceeded. You may need to change your search strategy. Helpful changes include reducing the number of queries, choosing a smaller database and setting an organism limit. You may also need to adjust the Algorithm parameters in the bottom section of the form. Choose a smaller number of target sequences, set a smaller Expect cut-off, use a larger word-size and turn on species specific repeats.

Job Title	NC_000011.10:c33892076-33858576 Homo sapiens...
RID	DVZE07CH013 Search expires on 07-03 01:39 am Download All
Program	Citation
Database	nr See details
Query ID	Id Query_19697
Description	NC_000011.10:c33892076-33858576 Homo sapiens chro...
Molecule type	dna
Query Length	33501
Other reports	?

Filter Results

Percent Identity	E value	Query Coverage
<input type="text"/> to <input type="text"/>	<input type="text"/> to <input type="text"/>	<input type="text"/> to <input type="text"/>
Filter		Reset

No significant similarity found. For reasons why, [click here](#)

Figure 10 : Protocol d'étude in silico de l'oncogène LMO2

3.1. Variants et protéine du LMO2(11p13) t(11;14)(p13;q11) :

- **Variant : t(7;11)(q35;p13) du LMO2(11p13) t(11;14)(p13;q11)**
- **Protéine : 4KFZ du LMO2(11p13) t(11;14)(p13;q11)**

QUERY: Full Text = "LMO2(11p13)" [Open In Query Builder](#) [JSON](#) [MyPDB Login](#)

Advanced Search Query Builder [Help](#)

Refinements [Clear All](#) [Q](#)

SCIENTIFIC NAME OF SOURCE ORGANISM [Clear](#)

Homo sapiens (5)
 Mus musculus (5)
 Drosophila melanogaster (1)

TAXONOMY [Clear](#)

Eukaryota (10)

EXPERIMENTAL METHOD [Clear](#)

SOLUTION NMR (5)
 X-RAY DIFFRACTION (5)

POLYMER ENTITY TYPE [Clear](#)

Summary Gallery Compact -- Tabular Report -- [↓ Score](#) [Download Files](#) All Selected

Display 25 per page

Displaying 1 to 10 of 10 Structures Page 1 of 1



4KFZ [Download File](#) [View File](#)

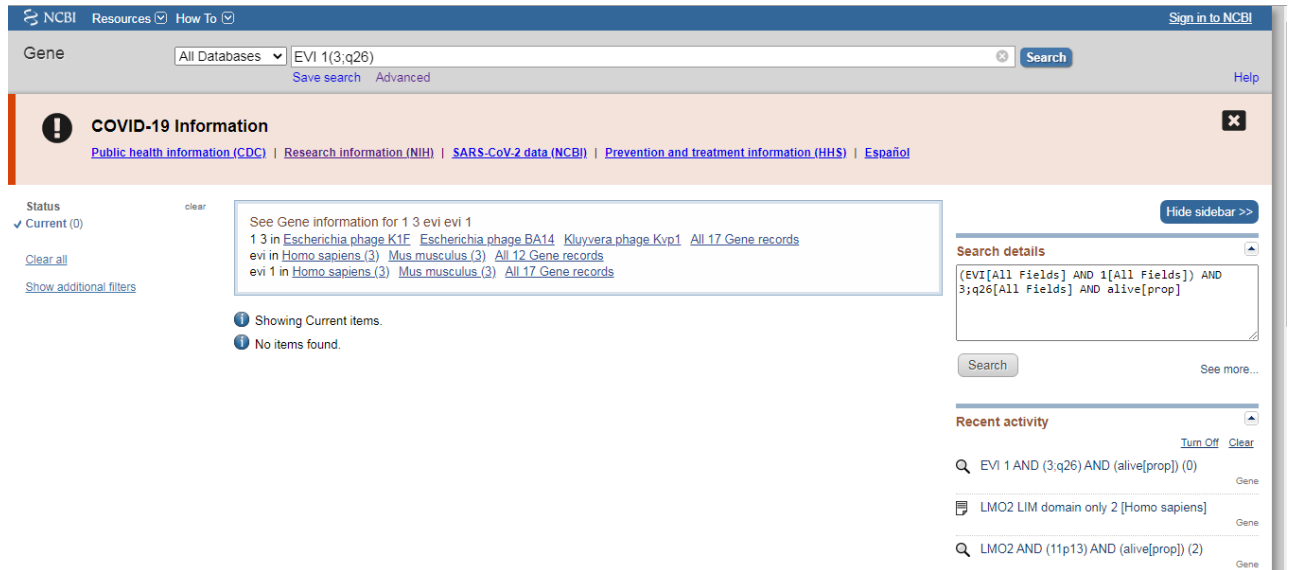
Crystal structure of LMO2 and anti-LMO2 VH complex

Sewell, H., Tanaka, T., El Omari, K., Cruz-Migoni, A., Mancini, E.J., Fuentes-Fernandez, N., Chambers, J., Rabbitts, T.H.

(2014) *Sci Rep* **4**: 3643-3643

Released 2014-01-22
Method X-RAY DIFFRACTION 2.8 Å
Organisms [Homo sapiens](#)
Macromolecule [Anti-LMO2 VH \(protein\)](#)
[LMO-2 \(protein\)](#)
Unique Ligands [ZN](#)

4. EVI 1(3;q26) t(3;3)(q21;q26) :



The screenshot shows the NCBI Gene database search results for the query "EVI 1(3;q26)". The search results are displayed in a table with columns for "Status", "Current (0)", and "Gene". The search details panel shows the query "(EVI[All Fields] AND 1[All Fields]) AND 3;q26[All Fields] AND alive[prop]". The recent activity panel shows the search history.

Status	Current (0)	Gene
✓	0	13 in Escherichia phage K1F Escherichia phage BA14 Kluyvera phage Kvp1 All 17 Gene records
		evi 1 in Homo sapiens (3) Mus musculus (3) All 12 Gene records
		evi 1 in Homo sapiens (3) Mus musculus (3) All 17 Gene records

Figure 11 : Protocol d'étude in silico de l'oncogène EVI1

- Variant : inv (3)(q21;q26) EVI 1(3;q26) t(3;3)(q21;q26)

5. LAZ 3/BCL6 (3q27) t(3;14)(q27;q32) :

NCBI Resources How To Sign in to NCBI

Gene LAZ 3/BCL6 (3q27) Search

Create RSS Save search Advanced Help

COVID-19 Information

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

Full Report ▾ Send to ▾ Hide sidebar >>

Showing Current items.

BCL6 BCL6 transcription repressor [*Homo sapiens* (human)]

Gene ID: 604, updated on 20-Jun-2021

[Download Datasets](#)

Summary

Official Symbol	BCL6 provided by HGNC
Official Full Name	BCL6 transcription repressor provided by HGNC
Primary source	HGNC:HGNC:1001
See related	Ensembl:ENSG00000113916 MIM:109565
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as	BCL5; LAZ3; BCL6A; ZNF51; ZBTB27
Summary	The protein encoded by this gene is a zinc finger transcription factor and contains an N-terminal POZ domain. This protein acts as a sequence-specific repressor of transcription, and has been shown to modulate the transcription of STAT-dependent IL-4 responses of B

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Expression
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from PubChem
- Interactions
- General gene information
 - Markers, Homology, Gene Ontology
- General protein information

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search

Advanced Help

COVID-19 Information

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

FASTA ▾ Send to ▾

Homo sapiens chromosome 3, GRCh38.p13 Primary Assembly

NCBI Reference Sequence: NC_000003.12

[GenBank](#) [Graphics](#)

>NC_000003.12:c187745472-187721377 Homo sapiens chromosome 3, GRCh38.p13 Primary Assembly

```
TCTTAGAACTGGGGTCTTAGAAGTGGTGATGCAAGAAGTTCTAGGAAAGGCCGGACACCAAGTGATTA
TTGCTGTTGCTGCGCCGCTGCTGCTGCTACTGCGCCGCGCCGCGCTGTGCGCGTGGTGGCGCTGCGCG
CGCCGCTGCTGATGATCATTATTTACCTTTTAATTTCTTTTTTTTCCGCTCTTGCCAAATGCTTTGGCT
CAAAGTTTCTATGTATCTATTGATATAAATGATATATTTATTTACTAGCTGTCAGGTGTTAAAA
TAAATGCCGAAGATTAGTCCACGCTCTCCCACCATAGGATATAGATTGTATGTTATTTATTATTATTA
TTGTTGCTTTGAGTGAATCGGCCGTTTGGGGAGGCTTTTGCCACCTCCCTTGTGTTGTTTGGTTTT
TGGAAAGGAGGTGGAGGAGAGGAGGGAATAGGGGGCGCCGAGGACAGAGGACGAGACAGTGC
TTGGGGGGTATCGGGCTAGTCTGGGGGGTCTGGCCCAAGACCGCGAGAGGACGCGCGCTCGCGCT
CTGCTCTTCTGCTGCTGCTTGGCTACGGCTTGTGTCTCTGGATTCTGCGGCTGTTTTTTCC
TCTTTCTCGCTGCAAACTGCTTCTCTGCTCGCTGCTGCTGGCCTCGCTTCCCTCCCTCCTCTTTT
TCCTCGCTCTTCTTCTTCCCTTCCCTGCTCTTGGGTAATCCGGGAGGCAAAAAGGAGGCTGCC
GCTCGCTGCGCCCGCTGGCTCTTCCATCGGCTCGGCTTTTTAGATTCCCTGCTTCCGAGCCGAACC
GAGATTTGGAAGGAAAAATGCAAGCAACTTTGGGTTTTTTGTTTTTTTTTTTGGTCATAAATCCA
GATGAGATGAAGTATCGTGTCTACTATTTCTTCTCAGAGCCGTGATCTTCTAATGAGAGCCGAGTGT
TGGTTTCCCTGCCCTGTGTTTTGGGGAGCTTGGGGGCGGGGATGTTGAACACTTTGAAAATTTCTT
TCTTGGTTGTAGTCGAGGAGAGTCGGGAACCTCCGCGCTCTCCCTGCTCATCCGAGGAGAGACTGA
TGGCGGGACAGGGCTTCTTGGGGTGGGTGGGAAAGGGGTGCGCACTGCACCTGCATGCTCTTCAAGTGTG
CGGGGGACTCAGGGAACTCGACAGGACAGCTTGGGAGAACGAGAAAGTGGCGGGATTTCTGCGCGCT
```

Change region shown

Whole sequence

Selected region

from: 187721377 to: 187745472

[Update View](#)

Customize view

Display options

Show reverse complement

Show gap features

[Update View](#)

Analyze this sequence

[Run BLAST](#)

[Pick Primers](#)

[Highlight Sequence Features](#)

BLAST® » blastx » results for RID-DW1HFKWB016 Home Recent Results Saved Strategies Help

[← Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

! There was a problem with the search. Please, contact [Help Desk](#) and include RID DW1HFKWB016. CPU usage limit was exceeded. You may need to change your search strategy. Helpful changes include reducing the number of queries, choosing a smaller database and setting an organism limit. You may also need to adjust the Algorithm parameters in the bottom section of the form. Choose a smaller number of target sequences, set a smaller Expect cut-off, use a larger word-size and turn on species specific repeats.

<p>Job Title NC_000003.12:c187745472-187721377 Homo sapiens...</p> <p>RID DW1HFKWB016 <small>Search expires on 07-03 02:15 am</small> Download All ▼</p> <p>Program Citation ▼</p> <p>Database nr See details ▼</p> <p>Query ID lcl Query_14588</p> <p>Description NC_000003.12:c187745472-187721377 Homo sapiens ch ...</p> <p>Molecule type dna</p> <p>Query Length 24096</p> <p>Other reports ?</p>	<p>Filter Results</p> <table style="width: 100%; border: 1px solid #ccc;"> <tr> <td style="text-align: center;">Percent Identity</td> <td style="text-align: center;">E value</td> <td style="text-align: center;">Query Coverage</td> </tr> <tr> <td style="text-align: center;"><input type="text"/> to <input type="text"/></td> <td style="text-align: center;"><input type="text"/> to <input type="text"/></td> <td style="text-align: center;"><input type="text"/> to <input type="text"/></td> </tr> <tr> <td colspan="2" style="text-align: center;">Filter</td> <td style="text-align: center;">Reset</td> </tr> </table>	Percent Identity	E value	Query Coverage	<input type="text"/> to <input type="text"/>	<input type="text"/> to <input type="text"/>	<input type="text"/> to <input type="text"/>	Filter		Reset
Percent Identity	E value	Query Coverage								
<input type="text"/> to <input type="text"/>	<input type="text"/> to <input type="text"/>	<input type="text"/> to <input type="text"/>								
Filter		Reset								

! No significant similarity found. For reasons why, [click here](#) [Feedback](#)

Figure 12 : Protocol d'étude in silico de l'oncogène LAZ3

5.1. Variants du LAZ 3/BCL6 (3q27) t(3;14)(q27;q32) :

- **(t(3;22)(q27;q11)**
- **t(3;4)(q27;p11)**

3.2 Discussion : Analyse & Discussion

Au vu des résultats bruts de l'interrogation des bases de données et du fait que nous avons plus d'informations sur les 03 premiers oncogènes , nous allons axer notre étude sur ceci , à savoir :

Les 03 Oncogènes objets de notre étude sont donc les suivants :

- **1. MYC (8q34)** t(8;14)(q24;q11) : **Variants (t(8;14)(q24;q32) , t(2;8)(p12;q24), t(8;22)(q24;q11))**
- **2. HOX11(10q24)** t(10;14)(q24;q11) : **Variant t(7;10)(q35;q24)**
- **3. LMO2(11p13)** t(11;14)(p13;q11) : **Variant t(7;11)(q35;p13)**

MYC (8q34) t(8;14)(q24;q11)

Ce gène est un proto-oncogène et code une phosphoprotéine nucléaire qui joue un rôle dans la progression du cycle cellulaire, l'apoptose et la transformation cellulaire. La protéine codée forme un hétérodimère avec le facteur de transcription apparenté MAX.

Ce complexe se lie à la séquence consensus d'ADN de la boîte E et régule la transcription de gènes cibles spécifiques. L'amplification de ce gène est fréquemment observée dans de nombreux cancers humains. Les translocations impliquant ce gène sont associées au lymphome de Burkitt et au myélome multiple chez les patients humains.

Il existe des preuves pour montrer que la traduction s'initie à la fois à partir d'un site de démarrage AUG en amont, dans la trame (CUG) et d'un site de démarrage AUG en aval, ce qui entraîne la production de deux isoformes avec des terminaisons N distinctes.

Expression omniprésente dans la vésicule biliaire (RPKM 49,6), l'œsophage (RPKM 44,6) et 25 autres tissus Une variété de translocations chromosomiques se produisent dans la leucémie aiguë lymphoblastique à cellules T (LAL-T) pédiatrique dans laquelle un oncogène cellulaire ou un gène lié à la croissance est transloqué vers le locus alpha/delta du gène du récepteur des cellules T.

Le t(8;14)(q24;q11) a été décrit au niveau cytogénétique et moléculaire, mais la maladie associée à cette translocation n'a pas été définie cliniquement. Quinze cas pédiatriques de leucémie/lymphome avec une translocation chromosomique t(8;14)(q24;q11) ont été recueillis à partir de publications antérieures et de dossiers institutionnels.

La prévalence estimée de cette anomalie parmi tous les cas de LAL était de 1%. La maladie t(8;14)(q24;q11) était caractérisée par une prédominance masculine (10/15), un âge médian de 5,5 ans (extrêmes 1,8-17 ans), un nombre élevé de globules blancs (médiane $95 \times 10^9 /l$), infiltration du système nerveux central (4/11), leucémie extramédullaire volumineuse (10/11) et immunophénotype des lymphocytes T (12/15). La survie médiane sans événement était de 4 mois et la survie médiane de 11 mois. Sept lignées cellulaires avec t(8;14)(q24;q11) ont été établies à partir de six des cas ; quatre étaient lymphoblastiques T, un était lymphoblastique T, mais exprimait des antigènes liés aux myéloïdes, et deux étaient à prédominance myéloïde. La leucémie/lymphome

t(8;14)(q24;q11) et d'autres LAL impliquant 13(q11) ont en commun une charge tumorale élevée, une propagation précoce aux sites extramédullaires, une propension à former des lignées cellulaires T-lymphoblastiques ou T-myéloïdes et, habituellement, une évolution clinique agressive.

Variants (t(8;14)(q24;q32) , t(2;8)(p12;q24), t(8;22)(q24;q11))

Le t(8;14)(q24.1;q32) et ses variantes – le t(2;8)(p12;q24.1) et t(8;22)(q24.1;q11.2) sont associés avec néoplasie à cellules B et entraîner un réarrangement du gène MYC/immunoglobuline (IG).

Nous avons corrélé les résultats cytogénétiques, moléculaires et clinico-pathologiques des patients présentant des translocations 8q24 vus dans le département d'hématologie, Christian Medical College, Vellore, de janvier 2003 à décembre 2015.

Il y avait 34 patients avec des translocations 8q24 (31, LAL et trois myélomes). Le t(8;14) a été observé chez 25 patients, t(8;22) chez sept et t(2;8) chez deux. Les principaux résultats étaient les suivants : 85 % d'hommes ; 79 % d'adultes, âge médian de 37 ans ; Morphologie L3 dans 61 % ; immunophénotype B mature dans 77 % ; maladie extra-médullaire dans 41 % ; anomalies supplémentaires chez 28 (85 %), notamment anomalies structurales des chromosomes 1q (41 %) et 13q (9 %) et monosomie 13 (15 %) ; caryotypes complexes dans 68%.

Il y avait deux lymphomes/leucémies à double hit, l'un avec at(14;18)(q32;q21) et l'autre avec at(3;14)(q27;q11.2), associés à un lymphome ganglionnaire à cellules B de haut grade et infiltrats leucémiques dermiques respectivement.

Seuls 13 échantillons ont été traités pour la PCR ADN et tous ces échantillons étaient positifs pour le réarrangement MYC-IgH (type c-gamma). Chez un seul patient, en plus du c-gamma, un réarrangement c-alpha a également été détecté.

La fréquence (1,7%) et la distribution de ces translocations dans notre série et l'association avec les anomalies 1q et 13q sont similaires à celles de la littérature. Les trisomies 7 et 12 ont été observées chez moins de 10 % de nos patients.

HOX11(10q24) t(10;14)(q24;q11) : Variant t(7;10)(q35;q24)

Ce gène code pour un facteur de transcription nucléaire qui appartient à la sous-famille des gènes à homéobox liés à NK ou NK-like (NKL). La protéine codée est nécessaire au développement normal de la rate au cours de l'embryogenèse. Cette protéine est également impliquée dans la spécification du destin des cellules neuronales. L'expression ectopique de ce gène due à des translocations chromosomiques est associée à certaines leucémies lymphoblastiques aiguës à cellules T. L'épissage alternatif donne lieu à plusieurs variantes de transcrits. [fourni par RefSeq, sept. 2010]

Expression biaisée dans la rate (RPKM 14.4) et les glandes salivaires (RPKM 3.2)

HOX11 est un gène codant pour une protéine homéoboîte qui est dérégulée dans la leucémie aiguë lymphoblastique à cellules T (T-ALL). Comme base pour étudier le mécanisme de dérégulation de l'expression de HOX11 dans la leucémie, le locus contenant le proto-oncogène HOX11 en 10q24 a été cloné à partir d'une banque de chromosomes artificiels P1 génomiques (PAC). Le clone PAC avec une taille d'insert de 120 kb a été isolé et cartographié par analyse de restriction.

Une série de sous-clones contigus ont ensuite été obtenus qui s'étendent sur 20 kb entourant le gène HOX11. Ces sous-clones ont été utilisés pour séquencer la totalité de la région de 20 kb jusqu'à la limite 3' de l'insert PAC. Ce travail fournit pour la première fois l'intron complet et les séquences non codantes 5' du gène HOX11, ce qui facilitera l'identification de nouveaux éléments de contrôle transcriptionnel pouvant être impliqués dans l'extinction de l'expression de HOX11 dans les cellules normales. Les informations de séquence ont également été utilisées pour rechercher de nouveaux grands cadres de lecture ouverts (ORF). Un tel ORF (1,1 kb) coderait pour une protéine d'au moins 39 kDa. Cette protéine basique (pI, 12,5) serait très riche en proline et pourrait potentiellement coder pour un nouveau facteur de transcription. Afin d'établir si cet ORF correspond à un gène transcrit de bonne foi, une analyse RT-PCR a été réalisée. L'ARNm de cette protéine est exprimé dans la lignée cellulaire T-ALL Jurkat et a été désigné HUG1, pour HOX11 Upstream Gene.

Une anomalie chromosomique courante dans la leucémie aiguë à lymphocytes T chez l'enfant est une translocation, t(10;14) (q24;q11), qui, avec la variante t(7;10)(q35;q24) est présente dans jusqu'à 7 % de ce type de tumeur. Le gène adjacent à la région 10q24 est transcriptionnellement activé après translocation soit en TCRD (14q11) soit en TCRB (7q35). Il code pour un gène homéoboîte étroitement lié aux gènes homéotiques régulés par le développement des mouches et des mammifères. La capacité de codage de ce gène activé, désigné HOX11, n'est pas perturbée dans une lignée cellulaire T portant la translocation t(7;10)(q35;q24). Par conséquent, le gène homéoboîte HOX11 semble être impliqué dans la tumorigenèse des cellules T.

LMO2(11p13) t(11;14)(p13;q11) : Variant t(7;11)(q35;p13)

LMO2 code pour une protéine à deux domaines LIM, riche en cystéine, nécessaire à l'érythropoïèse du sac vitellin. La protéine LMO2 a un rôle central et crucial dans le développement hématopoïétique et est hautement conservée. Le site de démarrage de la transcription LMO2 est situé à environ 25 kb en aval du cluster de translocation des cellules T 11p13 (11p13 ttc), où se produisent un certain nombre de translocations spécifiques de la leucémie lymphoblastique aiguë à cellules T. L'épissage alternatif donne lieu à plusieurs variantes de transcrits codant pour différentes isoformes.

Large expression dans la moelle osseuse (RPKM 14.3), les ganglions lymphatiques (RPKM 11.4) et 24 autres tissus

Leucémie aiguë lymphoblastique/lymphome lymphoblastique

Les translocations chromosomiques sont des événements primaires dans le développement des leucémies, représentant au moins une caractéristique génétique de la cellule souche cancéreuse putative. Les études de gènes influencés par les translocations chromosomiques ont fourni une grande quantité d'informations sur la façon dont le cancer est initié et maintenu. En particulier, les leucémies aiguës ont démontré que les translocations chromosomiques impliquent souvent des régulateurs de transcription qui fonctionnent en interagissant avec des protéines et en contrôlant le destin cellulaire dans le cadre aberrant de la cellule cancéreuse en développement. En tant que produit génique de translocation chromosomique par excellence, LMO2 possède de nombreuses propriétés qui caractérisent cette classe de molécules.

En plus de son implication dans les translocations chromosomiques, le gène LMO2 a été activé par inadvertance dans un essai de thérapie génique X-SCID par insertion rétrovirale. De nouvelles thérapies moléculaires ciblées directement sur la protéine LMO2 pourraient avoir un impact majeur en tant que compléments aux thérapies existantes ou en tant que thérapies à part entière. Dans cette revue, nous décrivons les connaissances actuelles sur le LMO2 et certaines voies possibles pour développer des réactifs qui pourraient être de possibles médicaments macromoléculaires à l'avenir.

Les leucémies sont des maladies clonales et acquises de la cellule souche hématopoïétique ou d'un précurseur déjà commis vers les lignées lymphoïde et/ou myéloïde. Leur classification repose sur des critères morphologiques, l'expression de marqueurs membranaires, et pour les leucémies

lymphoïdes, l'étude des réarrangements des gènes du récepteur T ou d'immunoglobulines qui permettent de caractériser le stade de différenciation des cellules leucémiques. Le clonage moléculaire de remaniements chromosomiques récurrents a permis d'identifier de nombreuses altérations géniques qui constituent des outils diagnostiques et de suivi thérapeutique maintenant pris en compte dans les nouveaux critères de classification des leucémies.

Le développement récent de thérapeutiques ciblées sur l'activité ou la stabilité d'une protéine oncogénique a été récemment illustré par les succès thérapeutiques obtenus dans le traitement des leucémies aiguës promyélocyaires et de la leucémie myéloïde chronique.

L'introduction en cytogénétique des techniques de bandes chromosomiques au début des années 1970 permit d'identifier précisément chaque chromosome et de caractériser des remaniements chromosomiques présents dans les cellules leucémiques. Trente-cinq pour cent des cellules leucémiques humaines ont un caryotype normal, alors que 65 % présentent un remaniement chromosomique le plus souvent unique.

Les anomalies les plus fréquentes sont principalement des inversions intra chromosomiques ou des translocations qui résultent de l'échange réciproque de matériel génétique entre deux chromosomes. Ces altérations sont acquises et présentes dans toutes les cellules leucémiques d'un même patient, ce qui témoigne de l'origine clonale de la leucémie. Leur caractère souvent récurrent, voire leur association systématique à certains types d'hémopathies suggéraient qu'elles pouvaient être l'événement causal du processus de transformation.

Il semblait donc possible que des oncogènes connus ou à découvrir étaient localisés à proximité des points de cassure chromosomique. Une étape fut franchie en 1982 lorsque les équipes de Croce et Leder démontrèrent que le gène *MYC*, identifié dans plusieurs modèles de leucémies lymphoïdes expérimentales et localisé chez l'homme en 8q24, était la cible des translocations du lymphome de Burkitt.

L'année suivante, il fut établi que la translocation t(9;22), caractéristique de la leucémie myéloïde chronique (LMC) et responsable de la formation du chromosome Philadelphie, impliquait un autre oncogène connu, ABL. Ces deux découvertes eurent un impact considérable et conduisirent à rechercher systématiquement des altérations chromosomiques dans les hémopathies malignes humaines.

Plus de 300 translocations récurrentes ont été rapportées. La liste de ces remaniements peut être consultée dans l'*Atlas of genetics and cytogenetics in oncology and haematology* (<http://www.infobiogen.fr/services/chromcancer>). La plupart des inversions et des translocations récurrentes ont été caractérisées au niveau moléculaire et ces études ont permis la découverte de nouveaux oncogènes. Les principaux remaniements chromosomiques conduisant à l'expression de protéines oncogéniques dans les leucémies.

D'une façon générale, les leucémies aiguës *de novo* s'accompagnent d'un blocage de différenciation et sont associées à des réarrangements chromosomiques impliquant des facteurs qui règlent l'activité transcriptionnelle. Deux types de gènes sont très souvent impliqués: des facteurs de transcription qui jouent un rôle majeur dans l'hématopoïèse normale et des gènes à homéodomaine ou des gènes qui règlent l'expression de gènes homéotiques.

Ces oncogènes peuvent bloquer la différenciation par plusieurs mécanismes: mutation de type perte de fonction, effet dominant négatif sur la copie résiduelle ou sur des facteurs transcriptionnelles hétérologues, recrutement de complexes co-activateurs ou corépresseurs de la transcription. Certains gènes sont altérés avec une fréquence particulière: le récepteur α de l'acide rétinoïque ou RAR α est remanié dans 5 translocations, 20 réarrangements altèrent les gènes du CBF (*core binding factor*) qui comporte deux sous-unités CBFA2 (ou AML1, RUNX1) et CBF β , et le gène *MLL*, qui est l'orthologue humain du gène *Trithorax* de la drosophile et qui règle positivement l'expression de gènes homéotiques, est impliqué dans plus de 40 translocations différentes.

Les oncogènes qui confèrent aux cellules un avantage de croissance ou de survie sans bloquer la différenciation sont plus généralement associés à des leucémies à cellules matures, aux lymphomes, ainsi qu'aux syndromes myéloprolifératifs comme la LMC et la leucémie myélomonocytaire chronique.

Conséquences des réarrangements chromosomiques : surexpression ou production d'une protéine de fusion

Expression anormale d'un gène de structure normale

Les réarrangements des gènes d'immunoglobuline (BCR) et du récepteur T (TCR) qui surviennent physiologiquement au cours du développement lymphoïde B et T et qui font intervenir une cassure double brin sur l'ADN par les recombinases RAG1 et RAG2, prédisposent à la

survenue de translocations par recombinaison illégitime. Ces translocations conduisent à l'expression aberrante d'oncogènes cellulaires, par un effet activateur de la transcription (mécanisme proche de la mutagenèse intentionnelle des leucémies expérimentales): l'oncogène est de structure normale mais un de ses deux allèles est sous le contrôle d'éléments activateurs présents dans les régions régulatrices des gènes du BCR ou du TCR.

Les oncogènes surexprimés dans les leucémies aiguës lymphoblastiques T (LAL T) codent le plus souvent pour des facteurs de transcription ou pour des protéines associées à des complexes transcriptionnels comme les protéines LYL1, TAL1/SCL, LMO1 et LMO2, HOX11 et HOX11L2 (*Figure 2*). Dans les hémopathies lymphoïdes à cellules plus matures, ce sont habituellement des gènes dont les produits règlent le cycle et la survie cellulaire qui sont surexprimés: le gène de la cycline D1 dans les lymphomes du manteau, *MYC* dans les lymphomes et leucémies de Burkitt et de rares LAL T, le gène antiapoptotique *BCL-2* dans les lymphomes folliculaires, alors que les gènes *TCL1* et *MTCP1* surexprimés dans les leucémies prolymphocytaires T favorisent la survie cellulaire du fait de leur association à la kinase AKT qui règle la survie cellulaire.

Dans les leucémies myéloïdes, le gène *EVII* est surexprimé dans de rares leucémies myéloblastiques aiguës (LAM) à la suite d'inversion $inv(3)(q21;q26)$ ou de translocations $t(3;3)(q21;q26)$. Ce gène avait été antérieurement identifié comme cible de mutagenèse insertionnelle dans des leucémies myéloblastiques chez la souris.

Production d'une protéine de fusion

Les remaniements chromosomiques conduisent souvent à la création d'un gène chimérique codant pour une protéine de fusion dont les propriétés sont différentes de celles de chacun des deux partenaires de la fusion. Plusieurs mécanismes peuvent participer aux propriétés oncogéniques de ces protéines: modification du niveau d'expression, de la localisation intracellulaire, de l'activité d'une protéine ou de son association à des protéines qui règlent son activité.

La production d'un gène de fusion est l'anomalie la plus fréquemment retrouvée dans les leucémies aiguës lymphoblastiques B (LAL B), les leucémies aiguës biphénotypiques de l'enfant et les leucémies myéloïdes qu'elles soient aiguës ou chroniques.

Certains gènes sont systématiquement et exclusivement remaniés dans les leucémies d'un même type. Ainsi, le gène *RARα* participe à la production de toutes les protéines de fusion associées aux leucémies aiguës promyélocyaires (LAP) et apparentées: PML-*RARα* pour la translocation t(15;17) la plus fréquente, PLZF-*RARα* pour la t(11;17), les fusions NPM-*RARα*, NuMA-*RARα* et STAT5B-*RARα* étant beaucoup plus rares. Ces observations suggèrent que l'interférence avec la fonction physiologique des rétinoïdes dans la différenciation myéloïde joue un rôle majeur dans le blocage de différenciation des LAP.

À l'inverse, le gène *AML1* est remanié dans des LAL B et des LAM. Ce gène, essentiel au développement de l'hématopoïèse primitive, code pour la sous-unité CBF α fixant l'ADN du complexe CBF.

La translocation la plus fréquente des LAL B de l'enfant, la t(12;21), conduit à la formation d'une protéine chimérique TEL-*AML1* contenant l'extrémité amino-terminale du facteur ETS TEL fusionné au domaine de liaison à l'ADN d'*AML1*, alors que, dans la translocation t(8;21) associée aux LAM, le domaine de fixation à l'ADN d'*AML1* est fusionné à la protéine ETO. La sous-unité CBF β , quant à elle, est l'un des deux partenaires du produit de fusion CBF β -SMMHC (chaîne lourde des chaînes de myosine du muscle lisse) résultant d'inversions du chromosome 16 associées à des LAM. Toutes ces protéines de fusion interfèrent avec la fonction d'*AML1* dans l'hématopoïèse.

La plupart des remaniements qui impliquent des protéines à activité tyrosine kinase conduisent à la formation d'un produit de fusion ayant une activité kinase constitutivement active. Dans les fusions impliquant un récepteur membranaire à activité tyrosine kinase (PDGFR dans certaines formes de leucémies myélomonocytaires chroniques, ALK dans les lymphomes anaplasiques à grandes cellules, FGFR1 dans les syndromes myéloprolifératifs atypiques avec anomalie de la bande chromosomique 8p12, les domaines extracellulaire et transmembranaire du récepteur sont absents de la fusion.

Le domaine cytoplasmique du récepteur, situé en position carboxy-terminale dans la protéine chimérique, est fusionné à un partenaire qui contient un motif d'oligomérisation nécessaire aux propriétés oncogéniques du produit de fusion. Ce type de motif est également et systématiquement présent dans le partenaire amino-terminal des protéines de fusion contenant des kinases intracellulaires (ABL, ARG [*Abelson related gene*], JAK2, SYK). Cela est en accord avec le mode

de fonctionnement de ces kinases qui sont physiologiquement activées en réponse à un stimulus extracellulaire par mémorisation et transphosphorylation.

La nature des séquences en amont de la kinase peut aussi influencer sur le type de la leucémie. Les leucémies associées à un produit de fusion BCR-ABL ont toutes pour origine une cellule immature multipotente, mais en fonction du point de cassure dans le gène *BCR*, les produits de fusion BCRABL sont associés à des LAL pro B (p190 BCR-ABL), à la LMC (p210 BCR-ABL) ou à une leucémie à polynucléaires neutrophiles (p230 BCR-ABL).

(ARGAGY, 2021)

Conclusion & Perspectives

Les leucémies humaines : un processus de transformation par étapes : Profils mutationnel variants oncogènes leucémiques par étapes de processus de cancérisation

Certaines leucémies se développent très tôt chez l'enfant. L'étude génétique de jumeaux identiques a montré que l'événement initiateur de la leucémie survenait *in utero*. La longueur et la variabilité de la latence des leucémies survenant chez des jumeaux ayant une translocation t(12;21) TEL-AML1 suggéraient que ce réarrangement favorisait mais n'était pas suffisant pour la transformation leucémique et que celle-ci dépendait de la survenue d'un événement additionnel. Cette hypothèse a été confortée par une étude systématique récente qui a mis en évidence des transcrits de fusion TEL-AML1 et AML1-ETO, par RT-PCR et par FISH, dans des cellules de sang de cordon avec une fréquence 100 fois plus élevée que la fréquence des leucémies associées à ces anomalies. La fréquence des cellules positives pour la fusion TEL-AML1 (10^{-3} à 10^{-4}) indique que la translocation survient dans une cellule immature qui prolifère mais se différencie en l'absence d'altération génétique additionnelle.

De même, il a été rapporté que la translocation t(14;18) associée aux lymphomes folliculaires était présente dans une sous-population de cellules lymphoïdes chez 50 % des sujets adultes et des transcrits de fusion BCR-ABL ont été décrits avec une fréquence de 30 % dans le sang périphérique d'adultes non leucémiques. Cependant, le très faible nombre de cellules positives suggère que ce réarrangement survient dans des cellules relativement matures et à capacité d'expansion limitée.

Altérations génétiques en l'absence d'anomalie cytogénétique

Des mutations ponctuelles et des microdélétions non détectables par les méthodes les plus performantes de cytogénétique, peuvent aussi conduire à l'activation d'oncogènes.

Des mutations ponctuelles activatrices des gènes *K-RAS* et *N-RAS* ont été décrites dans des syndromes myélodysplasiques et des LAM. Plus récemment, des mutations activatrices des récepteurs à activité tyrosine kinase FLT3 et c-KIT ont été mises en évidence avec une fréquence respectivement de 20-25 % et 5 % dans des LAM ayant une translocation impliquant les gènes du CBF, *MLL*, ou les oncogènes *PML-RAR α* et *DEK-CAN* (\rightarrow).

Des mutations de gènes de facteurs de transcription jouant un rôle majeur dans l'hématopoïèse normale ont également été récemment identifiées dans les leucémies. Ainsi, des

mutations du gène *C/EBP α* , qui code pour un facteur de transcription essentiel à la différenciation granulocytaire ont été détectées dans certaines LAM à caryotype normal et dans des myélodysplasies.

Nouvelles perspectives thérapeutiques

Différents modèles ont montré que l'avantage conféré par un oncogène rend les cellules tumorales dépendantes du maintien de l'activité et/ou de l'expression de la protéine oncogénique.

En effet, des cellules présentant une activation constitutive d'une voie de signalisation sont « dépendantes » de cette voie et plus sensibles que les cellules normales à son inhibition. Ainsi, des cellules tumorales ayant une activation constitutive de la kinase FRAP/mTOR à la suite de la délétion du gène suppresseur de tumeur *PTEN* sont plus sensibles que des cellules normales à un inhibiteur spécifique de mTOR. De même, il a été observé que les cellules de myélome arrêtent de proliférer et meurent par apoptose en présence d'inhibiteurs de protéasome qui inhibent l'activation de la voie NF- κ B. Dans des modèles de souris transgéniques pour un oncogène (*BCR-ABL* chez des souris développant une LAL B, ou *MYC* chez des souris développant des lymphomes T et une LAM) dont l'expression peut être réglée par la tétracycline, il a été montré que les masses tumorales régressent de façon spectaculaire lorsque le transgène cesse d'être exprimé.

Des résultats similaires ont été obtenus dans des modèles de tumeurs cutanées, du pancréas, de sarcome ostéogénique et de mélanome se développant chez des souris transgéniques pour une construction inductible de *MYC* ou *H-RAS* muté. La réversion de tumeurs à des stades avancés et dans lesquelles plusieurs événements oncogéniques coexistent (ce phénomène étant attesté par la clonalité des tumeurs hématopoïétiques) souligne l'intérêt du développement à des fins thérapeutiques d'agents pharmacologiques ciblant la fonction et/ou la stabilité de protéines oncogéniques.

De telles thérapeutiques ont déjà montré leur efficacité pour éliminer des cellules leucémiques. Le STI 571, un inhibiteur des kinases ABL, KIT et PDGFR, permet la mise en rémission de LMC et la régression de certaines tumeurs gastriques (GIST, *gastrointestinal stromal tumors*) dans lesquelles *KIT* est muté. La découverte de mutations de *FLT3* dans les LAM, et de *KIT* dans les LAM associées ou non à une mastocytose systémique permet d'envisager que les inhibiteurs spécifiques des formes mutées des kinases FLT3 et KIT qui ont été récemment développés pourraient être associés au traitement de certaines LAM. En raison de phénomènes de

résistance qui surviennent habituellement sous traitement par un seul inhibiteur, il est nécessaire de développer des modèles animaux reproduisant les leucémies humaines (par transgénèse ou par transplantation de moelle infectée par un virus contenant un oncogène) pour évaluer l'efficacité d'associations médicamenteuses. Une telle stratégie a été utilisée pour démontrer l'efficacité de l'association acide rétinoïque/arsenic qui conduit à la différenciation et à l'apoptose des cellules de LAP en interférant avec la fonction de PML-RAR α et en induisant la dégradation de cette protéine oncogénique.

Ainsi , en restant sur le principe de tendre vers la médecine prédictive et les recherches in silico en amont , ainsi en vue de RETRACER L'ÉVOLUTION d'une leucémie ; **La généalogie d'une tumeur ou l'arbre phylogénétique des leucémies** : Nous ouvrons cette parenthèse pour la suite de notre étude en vue d'établir un lien de parenté entre les différents oncog-nes et variants sur les leucémies.

En effet , dans un cancer, le principal danger vient des métastases de la tumeur primaire. La détermination de l'arbre de parenté des tissus cancéreux chez un même patient a montré que ces tumeurs secondaires ont pour origine plusieurs mutations clés survenues très tôt dans le foyer initial.

En règle générale, les mutations qui amènent les cellules à violer ce contrat de colocation et à proliférer de façon chaotique (les marques d'une tumeur maligne) sont éliminées. Les cellules mutées détectent leurs propres dysfonctionnements et se suicident (mort cellulaire programmée ou apoptose) ou sont détruites par le système immunitaire avant d'avoir endommagé l'organisme.

Toutefois, il arrive que s'accumulent des mutations contre lesquelles le système de surveillance cellulaire se révèle impuissant. Alors les tumeurs se développent et se propagent – comme le feraient des espèces invasives en milieu naturel. Les chercheurs connaissent certaines mutations à l'origine de la tumorigenèse, la formation de la tumeur initiale. Mais les véritables ennemis contre lesquels ils luttent sont les métastases. Celles-ci confèrent au cancer son caractère particulièrement létal. Ce sont des populations de cellules cancéreuses ayant migré à partir d'un foyer d'origine. Jusqu'à peu, la majorité des oncologues considéraient que les métastases apparaissaient à cause de nouvelles mutations dans les cellules tumorales et que ces mutations se produisaient relativement tard dans l'histoire de la tumeur primitive. Partant de ce principe, ils cherchaient à identifier ces mutations et à les cibler avec des médicaments.

Vers 2010, cependant, les avancées techniques ont permis aux scientifiques de séquencer à faible coût le génome humain tout entier (c'est-à-dire d'y déterminer l'ordre des constituants de l'ADN, les bases). Des groupes de recherche de plusieurs institutions ont alors commencé à étudier de manière exhaustive l'ADN des cellules tumorales. Mais à leur grande consternation, les chercheurs ont découvert qu'au sein même d'un unique patient, les tissus atteints présentaient des profils génétiques extrêmement variés. Agir sur toutes ces mutations pour soigner le cancer devenait dès lors difficile.

(ARGAGY, 2021)

Références bibliographiques

1. Oncogènes et leucémies: historique et perspective Oncogenes in leukemia: historical background and therapeutic prospect, Sylvie Gisselbrecht
M/S : médecine sciences, Volume 19, Number 2, février 2003
URI: <https://id.erudit.org/iderudit/000689ar> (Sylvie Gisselbrecht,2003)
2. Bioinformatique en oncologie: une discipline incontournable
Drs TIMOTHÉE OLIVIER, PIERRE O. CHAPPUIS et PETROS TSANTOULIS Rev Med Suisse 2016 ; 12 : 978-81
(Drs TIMOTHÉE OLIVIER & All, 2016)
3. Genome analysis Advance Access publication December 24, 2012 In silico identification of oncogenic potential of fyn-related kinase in hepatocellular carcinoma Jia-Shing Chen,y, Wei-Shiang Hung,y, Hsiang-Han Chan, Shaw-Jenq Tsai and H. Sunny Sun Vol. 29 no. 4 2013, pages 420–427 doi:10.1093/bioinformatics/bts715
(Jia-Shing Chen,y and All, 2013)
4. Charlotte Bouard. Étude in silico et caractérisation fonctionnelle des complexes de TWIST1. Cancer. Université de Lyon, 2016.
(Charlotte Bouard, 2016)
5. Bioinformatics | Oxford Academic , Volume 29, Issue 4 15 February 2013
(Oxford Academic, 2013)
6. ARCAGY-GINECO, 2021
(ARGAGY, 2021)
www.arcagy.org
7. Acute myeloid leukaemia Nicholas J Short, Michael E Rytting, Jorge E Cortes Lancet 2018; 392: 593–606 Published Online August 2, 2018 [http://dx.doi.org/10.1016/S0140-6736\(18\)31041-9](http://dx.doi.org/10.1016/S0140-6736(18)31041-9) Department of Leukemia (N J Short MD, Prof M E Rytting MD, Prof J E Cortes MD) and

Department of Pediatrics Patient Care (M E Rytting), The University of Texas MD Anderson Cancer Center, Houston, TX, USA

(Nicholas J Short and all 2018)

8. Acute Myeloid Leukemia Stem Cells: The Challenges of Phenotypic Heterogeneity
Marlon Arnone 1, Martina Konantz 1, Pauline Hanns 1, Anna M. Paczulla Stanger 2, Sarah Bertels 2, Parimala Sonika Godavarthy 2, Maximilian Christopeit 2 and Claudia Lengerke 1,2,*
(Marlon Arnone and All, 2020)

9. MYELODYSPLASTIC SYNDROMES Genetics of MDS Seishi Ogawa Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; and Department of Medicine, Center for Hematology and Regenerative Medicine, Karolinska Institute, Stockholm, Sweden © 2019 by The American Society of Hematology
blood® 7 MARCH 2019 | VOLUME 133, NUMBER 10
(Seishi Ogawa, 2019)

10. Dynamics of clonal evolution in myelodysplastic syndromes. Nat Genet. 2017.
(Makishima H, Yoshizato T, Yoshida K, et al)

11. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;
(Dohner H, Estey E, Grimwade D, et al.)

12. An update on classification, genetics, and clinical approach to mixed phenotype acute leukemia (MPAL) Maliha Khan¹ & Rabbia Siddiqi² & Kiran Naqvi¹ Received: 11 February 2018 /Accepted: 5 March 2018 /Published online: 15 March 2018 # Springer-Verlag GmbH Germany, part of Springer Nature 2018 Annals of Hematology (2018) 97:945–953 <https://doi.org/10.1007/s00277-018-3297-6>
(Maliha and All, 2018)

13. Genbank : GenBank ® is the NIH genetic sequence database
<https://www.ncbi.nlm.nih.gov/genbank/>

14. Ncbi : Le National Center for Biotechnology Information

<https://www.ncbi.nlm.nih.gov/>

15. Nih : National institutes of health

<https://covid19.nih.gov/>

16. WebBlast / Fasta : Basic Local Alignment Search Tool

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

17. Liens utiles à titre informatif (Plateformes & autres) :

- France Génomique :

www.france-genomique.org

- Listes des plateformes :

<https://www.france-genomique.org/plateformes-et-equipements/liste-plateformes/>

- Institut Imagine

<https://www.institutimagine.org/fr>

- Lyon Bio Pole

<https://lyonbiopole.com/>

- BioXclusters :

bioxclusters.wordpress.com

- **Sequençage génomique**

- 1ÈRE SOCIÉTÉ SAVANTE DÉDIÉE À LA MÉDECINE PERSONNALISÉE

<https://www.sfmpp.org/>

- Collectis : Entreprise d'ingénierie du génome spécialisée dans le développement d'immunothérapies fondées sur les cellules T CAR ingénierées

<https://www.collectis.com/research/gene-editing/>

- Auragen

<https://www.auragen.fr/>

- Sequoia

<https://laboratoire-sequoia.fr/>

