



Génétique et NAFLD chez l'Homme

Clémence CANIVET

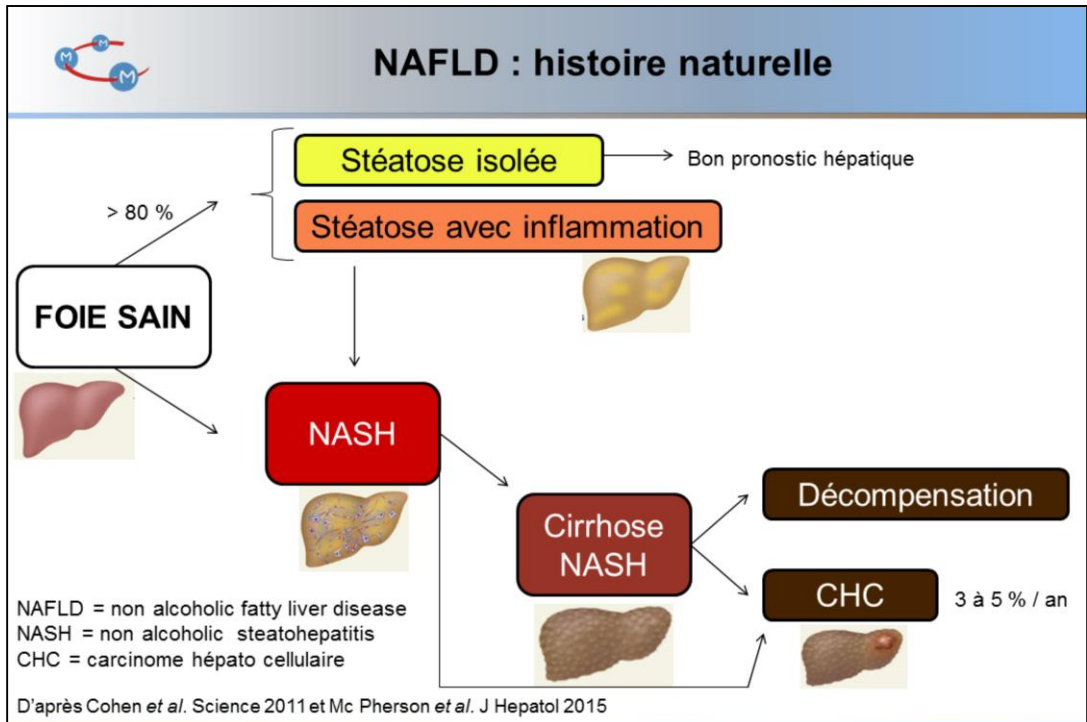
Interne en hépatologie

Master 2 - INSERM U1065 « complications hépatiques de l'obésité » C3M
CHU de Nice



Centre Méditerranéen de Médecine Moléculaire Inserm U1065

Mon sujet pour ce concours des internes porte sur la génétique et les maladies du foie gras non alcooliques chez l'Homme adulte.



Les maladies non alcooliques du foie (NAFLD) regroupent un ensemble d'entités allant de la stéatose (accumulation de triglycéries dans le foie), à la cirrhose en passant par la stéatohépatite (NASH = inflammation du foie gras).

Les NAFLD sont des maladies très fréquentes. On estime qu'environ 30% de la population américaine aurait une stéatose et 5% une NASH. Une NAFLD serait responsable de 60 à 70% des « perturbations inexplicées » du bilan hépatique.

Un patient peut évoluer en stéatose puis NASH mais peut aussi avoir directement une NASH sans passer par l'étape stéatose. Il existe 2 types de stéatose : la stéatose pure qui est de bon pronostic hépatique et qui n'évolue pas, et la stéatose avec inflammation, qui peut évoluer vers la NASH.

La présence d'une NASH est à risque d'évolution hépatique. L'évolution peut se faire classiquement vers la cirrhose puis le carcinome hépatocellulaire (CHC) mais le CHC peut apparaître avant le stade de cirrhose dans environ 30% des cas.



Quelques définitions

Polymorphisme : variations non pathogènes du génome qui sont la base de la diversité entre les individus (fréquence $>1\%$ dans la population). Un polymorphisme est une mutation et peut se situer en région codante ou non codante, sa longueur est variable.

Single nucleotide polymorphism (SNP) : polymorphisme de substitution au niveau d'un seul nucléotide. Les SNPs sont très nombreux ($>10^7$ par génome humain) et répartis sur tout le génome (1 tous les 300 bp)

NB : il existe plusieurs populations dites « ancestrales » (européenne, africaine, hispanique...). La répartition des polymorphismes est différente selon ces populations.

Centre Méditerranéen de Médecine Moléculaire Inserm U11065



Techniques d'identification des SNP

Etudes d'association entre un allèle donné d'un SNP et le phénotype « malade » ou « non malade »

Technique du gène-candidat

- Gènes connus
- Etude des gènes ayant un intérêt biologique => hypothèse à priori
- Problèmes :
 - Petit nombre de gènes d'intérêt
 - Faible puissance statistique des études

Technique de génome entier (GWAS)

- Gènes connus ou inconnus
- Etude du génome entier sans à priori
- Avantages :
 - Grand nombre de gènes étudiés
 - Grand effectif

Centre Méditerranéen de Médecine Moléculaire Inserm U11065

Le projet HapMap a permis de faire un séquençage du génome entier dans le début des années 2000.

Les techniques d'identification des SNP ont pour principe d'étudier l'association entre un allèle donné d'un SNP et le phénotype « malade » ou « non malade »

2 techniques existent :

- **La technique du gène candidat.** Elle repose sur des hypothèses à priori en étudiant un gène connu dont le produit d'expression est potentiellement impliqué dans le processus physiopathologique étudié. Les inconvénients de cette technique sont : le nombre limité de gènes à étudier, les cohortes ont des petits effectifs et donc la puissance statistique est faible.
- **La technique de génome entier (Genome Wide Association Study).** C'est l'étude de choix pour étudier les SNPs. Elle peut se faire par puce à ADN ou séquençage. Elle a l'avantage d'étudier l'ensemble du génome et donc les gènes connus et inconnus sans à priori. Cette technique d'identification des SNPs est indépendante de la biologie « traditionnelle ». Elle permet d'étudier un grand nombre de gènes à la fois et peut se faire sur des effectifs plus importants. Les risques de cette technique sont de mettre en évidence des faux positifs.

C'est pour cette raison que les études de GWAS doivent avoir des « p » beaucoup plus petits et les résultats doivent être validés sur une cohorte de validation de manière systématique.



NAFLD : une maladie « complexe »

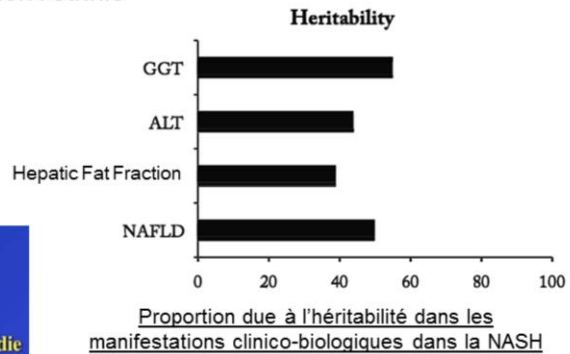
Polygénique : (30 à 50%)

- Agrégation familiale
- Différence de susceptibilité selon l'ethnie
- Polymorphismes génétiques

Environnement :

- Alimentation
- Sédentarité
- Obésité abdominale

Microbiote intestinal



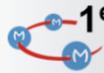
Don Giovanni et al. Curr Pharm Design 2013 ; Anstee et Day Nat Rev. Gastroenterol. Hepatol. 2015

Les NAFLD sont dites des maladies complexes : elles possèdent un déterminisme manifestement génétique mais dont la transmission ne correspond à aucun des modes mendéliens classiques de transmission, ni à un mode de transmission mitochondrial. Le déterminisme génétique est polygénique mais n'est pas suffisant pour expliquer la survenue de la maladie.

Le poids des gènes serait responsable de 30 à 50 % des NAFLD. Cela a été mis en évidence car il existe des agrégations familiales, des différences de susceptibilité selon l'ethnie et des polymorphismes génétiques favorisant/protecteur ont été découverts comme nous allons le voir dans les diapositives suivantes.

L'environnement joue également un rôle majeur dans le développement des NAFLD avec l'alimentation et la sédentarité entraînant une prise de poids, une obésité abdominale et la survenue d'un syndrome métabolique.

Enfin de manière plus récente, le microbiote intestinal jouerait aussi un rôle dans le développement des NAFLD.



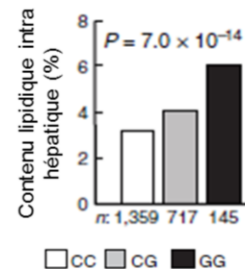
1^{er} SNP identifié : Patatine like phospholipase domain-containing 3 (PNPLA3)

Rs 738409 [C > G] : I148M

Association :

- ✓ à la **stéatose** ($p=7,0 \times 10^{-14}$)¹⁻²
- ✓ à la **stéatohépatite** OR = 3,48 ($p < 2 \times 10^{-4}$)³
- ✓ à la **fibrose** avec OR = 3,2 ($p < 1 \times 10^{-9}$)³

sans association avec l'insulino-résistance



Influence du PNPLA3 chez 2218 patients avec NAFLD diagnostiquée par spectro-IRM¹

¹ Romeo *et al.* Nat genet 2008 ; ² Speliotes *et al.* Hepatology 2010 ; ³ Sookoian *et al.* Hepatology 2011

La première étude GWAS dans les NAFLD a été publiée en 2008 par Romeo S *et al.* (Nat génétic 2008). Les patients, d'origine africaine, hispanique ou européenne présentaient des NAFLD diagnostiquées par spectro-IRM. La modification d'une cytosine par une guanine, entraînant en position 148 le changement d'une isoleucine en methionine (I148M) sur le gène de PNPLA3 (Patine like phospholipase domain-containing 3), était associée de manière très significative avec la stéatose ($p=5,9 \times 10^{-10}$). Ce SNP était retrouvé de manière plus importante dans la population d'origine hispanique.

Une meta-analyse en 2011 (Sookoian *et al.* Hépatology 2011) a montré que ce SNP était associé à la stéatohépatite avec OR à 3,48. Cette méta analyse a inclus 2124 sujets qui avaient bénéficié d'une biopsie hépatique.

Ce SNP était également associé à la fibrose. Récemment, Angula P *et al.* (Gastroenterology 2015) ont montré que la fibrose était la caractéristique histologique la plus associée à la mortalité dans les NAFLD.

De manière assez surprenante, ce SNP n'était pas associée avec l'insulino-résistance ni au diabète.

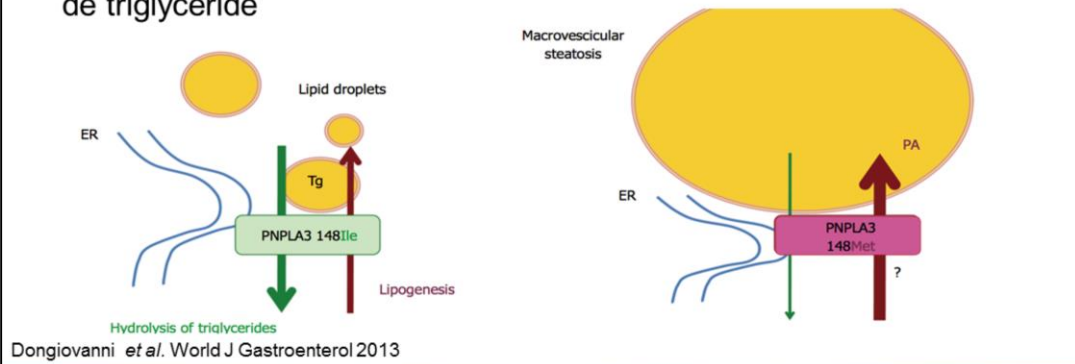


Rôle physiopathologique de PNPLA3

Protéine transmembranaire au niveau de la gouttelette lipidique dans les hépatocytes

Rôle dans le catabolisme des triglycérides

I148M => Modification proche du site catalytique => accumulation de triglycéride



PNPLA 3, également appelée adiponutrén, est une protéine de 481 acides aminés avec un poids moléculaire d'environ 53 kDa. Son gène est situé sur le chromosome 21. Cette protéine transmembranaire est présente au niveau de la gouttelette lipidique et au niveau du réticulum endoplasmique dans les hépatocytes.

Elle a pour rôle d'hydrolyser les triglycérides.

Dans le cas du SNP I148M, le changement de l'acide aminé se fait à proximité du site catalytique de la protéine. Cela entraîne un changement de conformation et une diminution de l'accès au site catalytique pour le substrat. La protéine mutée a des capacités diminuées pour hydrolyser les triglycérides et donc il va y avoir une augmentation de la taille des gouttelettes lipidiques et majoration de la stéatose macrovacuolaire.



Transmembrane 6 superfamily 2 (TM6SF2)

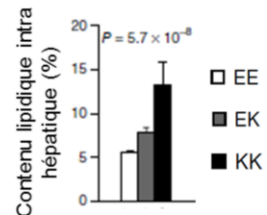
Rs 58542926 [C > T] : E167K

Association de l'allèle T :

- ✓ à la **stéatose** $p=5,7 \times 10^{-81}$
- ✓ à la **stéatohépatite** OR=1,84 [95%IC (1.23-2.79)]²
- ✓ à la **fibrose** $\beta=0.357 \pm 0.079$, [95%IC (0.203-0.511)]³

NB : l'allèle C est associé au métabolisme lipidique :

- ✓ avec le cholestérol total -8.4 +/- 1.6 mg/dL $p=7,6 \times 10^{-6}$ ⁴
- ✓ avec le HDL-cholestérol -3.7 +/- 0.9 mg/dL $p=3,2 \times 10^{-5}$ ⁴
- ✓ avec les triglycérides -9.4 +/- 2.1 mg/dL $p=9,5 \times 10^{-6}$ ⁴



Influence du TM6SF2 chez 2726 patients avec NAFLD diagnostiquée par spectro-IRM¹

¹ Kozlitina *et al.* Nat genetic 2014 ; ² Dongiovanni *et al.* Hepatology 2015 ; ³ Liu *et al.* Nat Commun 2014 ; ⁴ Pirola *et al.* Hepatology 2015.

En 2014, un nouveau SNP a été découvert dans les NAFLD. Il s'agit de TM6SF2 rs 58542926. Ce polymorphisme est caractérisé par le changement d'un acide aminé glutamine (E) par lysine (K) en position 167.

Ce polymorphisme est plus fréquent dans la population ancestrale européenne (7,2%) que chez les africains américains (3,4%) et les hispaniques (4,7%).

Kozlitina J *et al.* (Nat genetic 2014) ont été les premiers à décrire ce SNP comme associé à la stéatose dans une cohorte de patients dont la NAFLD était diagnostiquée par spectro-IRM.

Ces données ont été validées sur une cohorte de 1074 patients européens ayant bénéficiés d'une biopsie hépatique (Liu *et al.* Nat commun 2014). TM6SF2 était associé au degré de stéatose ($\beta=0,192 \pm 0,056$) [IC 95% (0,082-0,301)] ainsi qu'à la fibrose. L'association à la stéatohépatite n'était retrouvée que dans la cohorte de découverte et non dans celle de validation dans cette étude.

Dongiovanni P *et al.* (hepatology 2015) ont également validé les données concernant la fibrose et ont trouvé une association avec la stéatohépatite dans une cohorte de 1200 patients avec une NAFLD

prouvée histologiquement.

De manière intéressante, la méta analyse de Pirola CJ *et al.* (Hepatology 2015) mettait en évidence le fait que l'allèle T était associé au spectre des complications métaboliques hépatiques tandis que l'allèle C était associé au métabolisme lipidique (cholestérol total, HDL et LCL cholestérol).

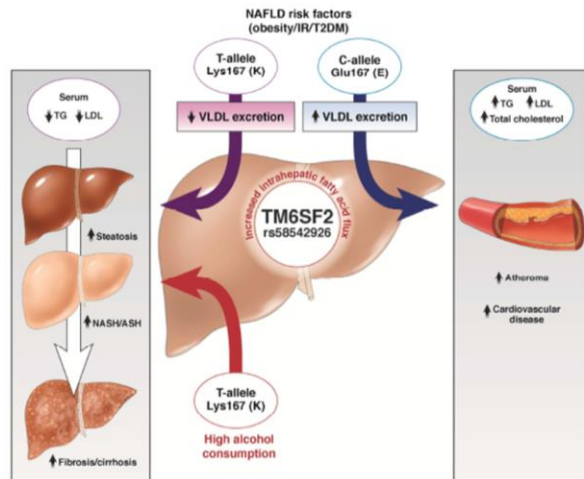


Rôle physiopathologique de TM6SF2

Protéine transmembranaire présente dans le réticulum endoplasmique du foie, de l'intestin grêle et du rein.

Fonction très mal connue :

- Transporteur lipidique ?
- Régulation du VLDL ?
- Modulateur de l'orientation des lipides entre le foie et les vaisseaux ?



Modèle hypothétique de la régulation des lipides par TM6SF2

Anstee *et al.* Gastroenterol 2016

Actuellement les données physiopathologiques concernant la protéine TM6SF2 sont très limitées. Il s'agit d'une protéine à 7 domaines transmembranaires située dans le réticulum endoplasmique du foie, de l'intestin grêle et du rein. Son gène est situé sur le chromosome 19.

Sa fonction n'est pas connue.

Anstee QM *et al.* (Gastroenterology 2016) dans une revue récente, ont proposé un modèle hypothétique de régulation des lipides par TM6SF2. En présence de l'allèle T (qui donne l'acide aminé lysine), il y aurait une diminution de l'excrétion hépatique de VLDL, ce qui serait à l'origine d'une accumulation de triglycérides dans le foie entraînant une stéatose puis une NASH. En revanche, en présence de l'allèle C (qui donne l'acide aminé glutamate), il y aurait une augmentation de l'excrétion du VLDL à l'origine d'une protection hépatique et d'une majoration des lipides circulants responsables d'athérome et de maladies cardiovasculaires.



Autres SNPs validés ou en cours de validation par plusieurs équipes

Nom du SNP	Stéatose	NASH	Fibrose
GCKR ¹ rs 780094 [C > T]			
NCAN ² rs 2228603 [C > T]			
PEMT ³ rs 7946 [G > A]			
FNDC5 ⁴ rs 3480 [G > A]			

¹ Petta et al. Plos One 2013 ; ² Gordon et al. Hum Hered 2013 ; ³ Tan et al. pharmacogenet genomics 2016 ; ⁴ Anty et al. J Hepatol EASL 2015

Beaucoup d'autres SNPs ont été décrits dans la littérature mais leur association aux NAFLD n'est pas claire et les résultats ne sont pas reproductibles d'une équipe à l'autre.

J'ai choisi les 3 SNPs qui avaient les meilleures preuves. Le dernier SNP a été découvert par notre équipe.

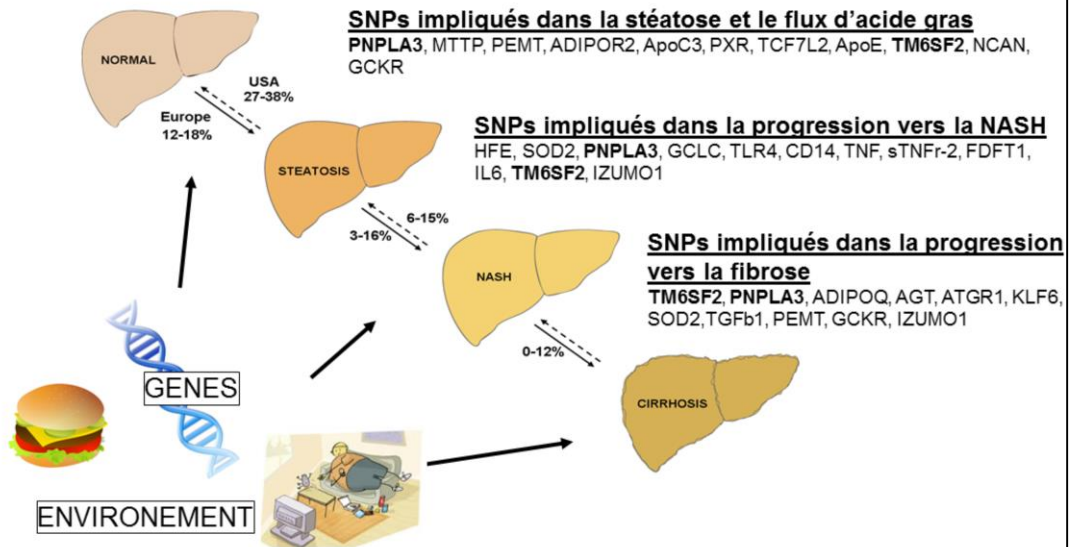
- **GCKR (glucokinase regulatory protein) rs780094.** Ce gène code pour une protéine produite par les hépatocytes, et est une enzyme clé du métabolisme glucidique. Ce SNP était associé à la fibrose > F1 (OR = 1,88 avec $p < 0,001$) dans une cohorte de 366 patients européens qui ont tous eu une biopsie hépatique. Ce SNP était également associé à l'élévation des triglycérides dans le sang ($p = 0,02$). Il n'a pas été étudié avec la stéatohépatite.
- **NCAN rs 2228603.** Ce locus contient au moins 20 gènes sur le chromosome 19. Aucun de ces gènes ne semblent jouer un rôle dans le métabolisme lipidique ou dans les NAFLD. NCAN code pour une protéine de la famille des protéoglycanes, la neurocan, qui a été initialement décrite dans le système nerveux et participe à l'adhésion cellulaire. Le SNP était associé à la stéatose ($p = 0,003$), à l'inflammation lobulaire ($p = 0,002$) et à la fibrose péri-sinusoidale ($p = 0,002$) dans une cohorte de 1092 patients obèses morbides ayant bénéficiés d'une biopsie hépatique au cours d'une chirurgie

bariatrique. Ce SNP est en faveur d'un axe cerveau-foie.

- **PEMT (Phosphatidylethanolamine N-methyltransferase) rs 7946.** Ce gène code pour une protéine du même nom qui catalyse la biosynthèse de la phosphatidylcholine, élément essentiel dans la formation et l'export des triglycérides sous forme de VLDL. Une diminution de PEMT entraîne une accumulation de triglycérides dans le cytoplasme des hépatocytes. Cette publication est une méta-analyse (6 articles) sur les NAFLD et PEMT. Le diagnostic de NAFLD était fait par histologie pour 2 études et par radiologie pour les 4 autres études. Ce SNP était associé à la NASH (OR = 3,76 avec $p < 0,00001$). Le diagnostic de NASH se basait sur la biopsie et la méta analyse ne prend en compte que les 2 études ayant réalisées des biopsies.
- **FNDC5 (Fibronectin type III domain-containing 5) rs 3480.** Ce gène code pour une myokine "l'irisine". Son rôle est de participer à la brunisation du tissu adipeux et à la libération d'énergie sous forme de chaleur. Il a été montré très récemment, qu'il s'agit également d'une hépatokine jouant un rôle dans le métabolisme glucidique. Le SNP découvert est protecteur de stéatose (OR = 0.69 avec $p = 0,004$) et de fibrose ($\beta = -0,23$ avec $p = 0,007$). C'est la première fois qu'un SNP protecteur est découvert dans les NAFLD. Les patients avaient eu une biopsies hépatique, soit de manière systématique au cours d'une chirurgie bariatrique (cohorte d'étude), soit pour anomalies du bilan hépatique chez des patients obèses (cohorte de validation).



Résumé des SNPs impliqués dans les NAFLD



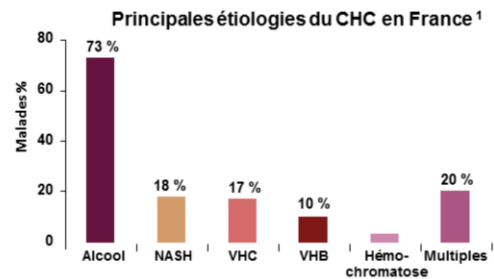
Voici un résumé des principaux SNPs décrit à l'heure actuelle dans les NAFLD.



SNPs impliqués dans le carcinome hépatocellulaire sur NAFLD

Actuellement les NAFLD sont la deuxième cause de carcinome hépatocellulaire (CHC) en France.

La particularité du CHC est qu'il survient dans 30% des cas sur foie non cirrhotique².



Nom du SNP	Risque	Référence
PNPLA3	OR=2,26 (p = 0,0082)	Liu YL <i>et al.</i> J Hepatol 2014 Trepo E <i>et al.</i> J Hepatol 2016
TM6SF2	Pas d'association retrouvée (p = 0,42)	Liu YL <i>et al.</i> Nat Commun 2014

¹ Rosa *et al.* Hepatology. 2010.1144A ; ² Mittal *et al.* Clin Gastroenterol Hepatol 2015

Certaines études ont également étudié la génétique dans les carcinomes hépatocellulaires (CHC) sur NAFLD.

A l'heure actuelle, les NAFLD sont la 2^{ème} cause de CHC après l'alcool.

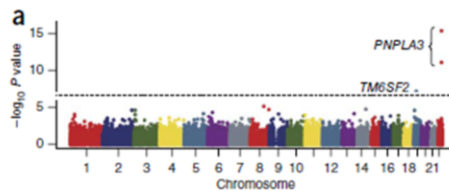
La particularité du CHC sur NAFLD est qu'il survient dans 1/3 des cas sur un foie non cirrhotique.

Seule PNPLA3 a été retrouvée comme associée au risque de CHC.



Take home messages

- Les NAFLD sont des maladies complexes
- La génétique a un poids important dans le développement de ces maladies
- PNPLA3 et TM6SF2 sont les mieux caractérisés
- PNPLA3 favorise la stéatose par une voie alternative à l'insulino-résistance



Centre Méditerranéen de Médecine Moléculaire Inserm U1065



Perspectives

- De nouveaux SNPs sont en cours de validation
- L'identification des SNPs permettrait au niveau :
 - physiopathologique, d'améliorer les connaissances dans les NAFLD
 - pronostique, d'identifier les patients à risque de complications vasculaires ou de cancers (CHC)
 - thérapeutique, d'adapter la prise en soins des patients avec NAFLD selon leurs particularités génétiques