



**HAL**  
open science

# Étude de l'hémoglobine glyquée chez des patients diabétiques de type 1

Emilie Gloaguen

► **To cite this version:**

Emilie Gloaguen. Étude de l'hémoglobine glyquée chez des patients diabétiques de type 1. Médecine humaine et pathologie. Université Paris Cité, 2020. Français. NNT : 2020UNIP5046 . tel-03615690

**HAL Id: tel-03615690**

**<https://theses.hal.science/tel-03615690>**

Submitted on 21 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Paris

**Ecole doctorale 393 – Pierre Louis de Santé Publique à Paris :**  
**Epidémiologie et Sciences de l'Information Biomédicale**  
***UMR-S958 Génétique du diabète***

# **Étude de l'hémoglobine glyquée chez des patients diabétiques de type 1**

Par Emilie GLOAGUEN

Thèse de doctorat d'épidémiologie génétique

Dirigée par Flavie MATHIEU

Présentée et soutenue publiquement le 3 Septembre 2020

Devant un jury composé de :

Marie-Aline CHARLES, DR, Université de Paris, Présidente du jury  
Martine VAXILLAIRE, DR, Université de Lille, Rapportrice  
Hervé PERDRY, MCU, Université Paris Saclay, Rapporteur  
Jean-Claude CAREL, PUPH, Université de Paris, Examineur  
Marc De Kerdanet, PH, CHU de Rennes, Examineur  
Flavie MATHIEU, IR, Université de Paris, Directrice de Thèse



Except where otherwise noted, this is work licensed under  
<https://creativecommons.org/licenses/by-nc-nd/3.0/fr/>





# Résumé en français

---

## Titre : Étude de l'hémoglobine glyquée chez des patients diabétiques de type 1

**Résumé :** Malgré les avancées technologiques au cours des 20 dernières années, le contrôle de la glycémie demeure un défi de taille pour les enfants atteints de diabète de type 1 (DT1), certains patients n'atteignant jamais un équilibre glycémique correct (estimé par le taux d'hémoglobine glyquée, HbA1c), ce qui entraîne des complications sévères à long terme. À ce jour, il n'existe aucune étude conjointe des facteurs cliniques, épidémiologiques et génétiques influençant le taux d'HbA1c. L'objectif de cette thèse est donc l'étude de l'ensemble de ces facteurs, en trois volets distincts.

Une première étape a consisté en l'étude des propriétés statistiques d'un test d'association génétique, le « General Regression Model » (GRM), dans le cadre de l'étude d'un trait quantitatif. En effet, les études d'association génétique à grande échelle d'un trait quantitatif utilisent le plus souvent des modèles de régression linéaire (LR) faisant l'hypothèse d'un effet dose de l'allèle à risque (modèle additif LR\_ADD), ce qui entraîne une perte de puissance en cas d'écart à l'additivité. Au contraire, le GRM ne fait pas d'hypothèse quant au mode de transmission et permet de le tester formellement. Nous avons comparé les propriétés statistiques du GRM à celles des tests LR pour un grand nombre de modèles génétiques par une étude de simulation. Les résultats obtenus montrent que le GRM peut être aussi puissant que le LR\_ADD sous un modèle additif et qu'il est plus puissant que le LR\_ADD en cas d'écart à l'additivité. Le test du modèle de transmission a permis de conclure au modèle simulé dans la plupart des cas.

Dans un deuxième temps, nous avons réalisé une revue systématique de la littérature des facteurs influençant le taux d'HbA1c chez le jeune patient DT1 entre 2010 et juillet 2016. Pour les facteurs non-génétiques, 1974 articles ont été répertoriés dont 78 ont été sélectionnés. La plupart de ces articles portent sur l'étude de l'effet d'un facteur particulier sur le taux d'HbA1c. Les principaux facteurs influençant le taux d'HbA1c sont : le statut pubertaire, le statut socio-économique, l'ethnie, la durée du DT1, l'adhérence et le type de traitement. Pour les facteurs génétiques, 277 ont été répertoriés et 14 sélectionnés. La plupart de ces études se sont intéressées à des gènes candidats dans des cohortes de petite taille. Ce travail souligne la nécessité d'étudier conjointement les différents facteurs influençant le taux d'HbA1c.

Enfin, en collaboration avec le Dr Bendelac et le Pr Nicolino, Directeur du service d'endocrinologie-diabétologie pédiatrique à l'Hôpital Femme-Mère-Enfant de Lyon, j'ai participé au projet GENDIR qui vise à étudier le rôle de facteurs impliqués dans le contrôle de l'équilibre glycémique. Cette étude rétrospective porte sur un échantillon de 362 patients pour lesquels nous avons collecté les principaux facteurs identifiés dans la revue après 2, 5 et 10 ans de diabète et géotypé 260 patients. L'étude préliminaire a permis d'identifier l'origine géographique des parents comme étant le principal facteur influençant le taux d'HbA1c après de 2 ans de diabète après correction pour la dose d'insuline requise. Aucun marqueur génétique n'a été répliqué par manque de puissance, et nous avons pu mettre en évidence différents profils d'évolution du taux d'HbA1c, qui devront être étudiés plus en détails dans la suite de l'étude.

Sur le long terme, l'ensemble des résultats obtenu dans le cadre de ma thèse pourrait avoir des applications directes en clinique et bénéficier aux personnes souffrant de diabète au moment du diagnostic (limitation de la période de titrage et d'ajustement des doses d'insuline), au cours de la maladie (définition de la dose d'insuline optimale de chacun des personnes souffrant de diabète en fonction de leurs caractéristiques), limitant ainsi les complications du diabète.

**Mots clefs :** Hémoglobine A glycosylée ; Équilibre glycémique ; Génétique ; Étude d'association pangénomique ; Modes de transmission héréditaire

# Résumé en anglais

---

**Title:** Study of the glycated hemoglobin in type 1 diabetes patients

**Abstract :** Despite technological advances over the past 20 years, the control of blood glucose level is still a major challenge for children with type 1 diabetes (T1D), with some patients never achieving correct glycemic control (estimated by glycated hemoglobin level, HbA1c), leading to severe long-term complications. To date, there is no joint study of the clinical, epidemiological and genetic factors influencing HbA1c level. The objective of my PhD was therefore to study all these factors and is organized in three parts.

First we studied the statistical properties of a genetic association test, the General Regression Model (GRM), for a quantitative trait. Indeed, large-scale genetic association studies of quantitative traits often use linear regression (LR) models assuming a dose effect of the risk allele (additive model, LR\_ADD), resulting in a loss of power in case of deviation from the additivity. The GRM does not make any assumptions about the mode of transmission and formally test it. We compared the statistical properties of GRM to LR tests for a large number of genetic models by a simulation study, showing that GRM may be as powerful as LR\_ADD under an additive model, and can be more powerful when there is a deviation from the additivity. The transmission model test concluded to the simulated model in most cases.

In a second step, we systematically reviewed the literature on factors influencing HbA1c level in young T1D patients between 2010 and July 2016. For non-genetic factors, 1974 articles were identified and then 78 selected. Most of these articles studied the effect of a particular factor on HbA1c levels. The main factors influencing HbA1c levels are pubertal status, socio-economic status, ethnicity, duration of T1D, adherence and treatment. For genetic factors, 277 articles were identified and 14 were selected. Most of these studies have focused on candidate genes in small samples. This work highlights the importance to jointly study factors influencing HbA1c.

Finally, I participated to the GENDIR project, which aims to study the role of factors involved in glycemic control in collaboration with Dr. Bendelac and Professor Nicolino, who is handling the paediatric endocrinology-diabetology department of the Femme-Mère-Enfant Hospital at Lyon. This retrospective study was performed on a sample of 362 T1D patients. We collected the main factors identified in the review after 2, 5 and 10 years of diabetes and 260 patients were genotyped. The preliminary study identified the geographical origin of the parents as the main factor influencing HbA1c levels after 2 years of T1D after correction for insulin dose required. No genetic marker was replicated, and we were able to highlight different patterns of evolution of HbA1c levels, which will need to be studied in more detail in the rest of the study.

The results of my PhD would have clinical applications and will benefit to diabetic patients at the time of diagnosis (limitation of the titration and adjustment period of insulin doses), during illness (definition of the optimal insulin dose for each diabetic patient according to their characteristics), thus limiting diabetes complications.

**Keywords:** Glycated Hemoglobin A; Glycemic equilibrium; Genetic; Genetic Association Studies; Inheritance Patterns

# Remerciements

---

Je tiens en premier lieu à remercier l'ensemble du jury. Je remercie Marie-Aline CHARLES d'avoir accepté la présidence de mon jury de thèse. Je remercie aussi Hervé PEDRY et Martine VAXILLAIRE, d'avoir accepté de relire mon manuscrit de thèse et d'être rapporteur de ma thèse. Et je remercie enfin Marc DE KERDANET et Jean-Claude CAREL de bien avoir accepté de participer au jury de ma thèse.

Ensuite, je me dois de remercier tout particulièrement Flavie MATHIEU, ma directrice de thèse, sans qui cette thèse n'aurait pas pu voir le jour. Je la remercie de m'avoir fait confiance pour mener à bien ce projet, malgré les nombreuses embuches et rebondissements qui ont eu lieu durant ces six dernières années. Ces années passées à travailler avec vous m'auront beaucoup appris, dans les domaines de la recherche et de la science, mais aussi de manière plus générale sur ma façon de travailler.

Je remercie aussi Cécile JULIER, pour m'avoir permis de commencer ma thèse dans son laboratoire, pour son soutien logistique et ses questionnements scientifiques, qui n'ont pu qu'améliorer cette thèse. Je remercie aussi l'ensemble de son laboratoire durant les trois premières années de cette thèse : Anne PHILIPPI et Anne DEGAVRE pour leur aide entre autre dans le domaine de l'informatique, domaine qui m'était largement inconnu avant de débiter cette thèse ; Valérie SENEÉ pour le soutien moral forcément nécessaire pour arriver au bout d'une thèse ; Claire VANDIEDONCK pour les précieux conseils qu'elle a su me prodiguer dans un très grand nombre de domaines pour que je puisse mener cette thèse à son terme.

Je remercie aussi les différentes personnes avec qui j'ai pu collaborer durant cette thèse. Dans un premier temps, je souhaite remercier le Dr Nathalie BENDELAC et le Pr Marc NICOLINO, pour leur aide sur la revue systématique de la littérature, mais aussi pour leur implication ainsi que celles des différents services des Hospices Civils de Lyon sur la cohorte GENDIR, et plus particulièrement les Dr Isabelle ROUVET et Pr Laurent Schaeffer de la Biothèque du Centre de Biotechnologie Cellulaire des Hospices Civils de Lyon. Je tiens aussi à remercier les différents financements qui ont permis projet GENDIR de voir le jour : Lilly, le Centre National de Génotypage et l'Aide aux Jeunes Diabétiques. Je souhaite aussi remercier les membres de l'unité UMR-8199 et de la cohorte D.E.S.I.R. avec qui nous avons pu échanger pour le projet GRM : Philippe FROGUEL et Beverley BALKAU, pour leur contribution et les relectures du papier, Mathilde BOISSEL, Ghislain ROCHELEAU et Mickaël CANOUIL pour les différentes réunions que nous avons fait sur les données D.E.S.I.R. et sur le GRM de manière générale qui n'ont fait qu'améliorer le papier.

Cette thèse n'aurait sans doute jamais pu se terminer sans l'aide de France MENTRE, directrice de l'équipe BIPID au sein de l'unité IAME UMR 1137 et du DEBRC, qui m'a permis de terminer cette thèse dans les meilleures conditions possibles. Merci de m'avoir recueillie et hébergée quelques mois avant de m'offrir la possibilité de travailler au sein du DEBRC. Je remercie aussi l'ensemble de l'équipe BIPID et du DEBRC de m'avoir si chaleureusement accueilli, et plus

particulièrement Jimmy MULLAERT, pour ses conseils et sa disponibilité pour m'aider sur cette thèse.

Je remercie aussi mes différents groupes d'amis de longue date, dont les deux Camille, Roxanne, Norine, Eve et Nhi ; Tamara et Lydie. Nos projets respectifs font qu'on se voit sans doute beaucoup moins que ce que l'on voudrait, mais l'important c'est qu'on prenne toujours le même plaisir à se retrouver. Merci d'avoir été présentes quand il le fallait, pour me changer les idées quand il le fallait, pour mieux se reconcentrer sur cette thèse après.

Cette thèse m'a aussi permis de créer des liens forts, et en premier lieu mes Copinollègues : Marine, Mathilde et Sophie. Quand j'y pense, nous n'avons pas passé tant de temps que ça au même endroit, mais nous avons pu y lier une amitié très forte. Merci de m'avoir soutenue et supportée pendant tout ce temps. Merci d'être tout le temps là quand il faut, et surtout le 4<sup>e</sup> vendredi de chaque mois !

Je remercie aussi la petite team de l'ENVA : Marie-Perle, Séléna, Mathilde et Jeanne. Puisque cette année-là a été un peu particulière pour moi, pleine de doute et de questionnement, sur ce que je voulais vraiment faire de ma vie, je vous remercie de l'avoir partagée avec moi. Je ne pourrais peut-être pas soigner ma Naïade toute seule, mais je saurais chez qui l'envoyer les yeux fermés ! Un jour, ce sera à votre tour d'écrire une thèse, et qui sait peut-être que l'une d'entre vous décidera même de devenir PhD en plus d'être VetD.

Ensuite, je remercie ceux qui ont fait naître ma passion pour la création de Doodle et de groupe Messenger (avec comme point d'origine le groupe *IAME & Cie*) pour des évènements qui finalement n'ont jamais lieu, soi-disant à cause de grèves ou de virus ; ceux qui sont au moins allés une fois à L'Escale ; ceux que je gave de gâteaux plus ou moins réussis. Pour certains d'entre vous, ça sera bientôt à votre tour d'écrire !

Je remercie en premier ce qui ne sont pas ou plus en thèse : Aurélien, Jinju, Manel, Marc, Marilou, Morgane, Nadège, et Saadia. Avec une petite place à part pour Coralie, ma compagne de bureau, au-delà du bureau 437 puisque je te retrouve maintenant à l'URC. Je te remercie pour les séances karaoké, ainsi que pour ton aide et tes précieux conseils à chaque fois que j'en avais besoin. Je suis sûre que d'ici un an ou deux, tu changeras de catégorie dans ces remerciements.

Je remercie ensuite ceux qui vont devoir passer de longues heures à écrire, certains avec passion, d'autre beaucoup moins : Antonio, merci d'avoir toujours été là quand je devais faire un graph en ggplot, mais surtout quand j'avais besoin de parler, de me plaindre de tout et de n'importe quoi ; Claire, merci pour ta joie de vivre communicative ; Drifa, mon autre compagne de bureau et de séance ciné ; François, je suis obligée de t'associer à L'Escale et aux trajets jusqu'à St-Lazare qui suivent, mais aussi à nos discussions dont celles sur nos thèses respectives, aux séances cinés annulées car complètes, je te remercie pour tout ça, et aussi pour tout un tas de choses que j'ai forcément oublié ; Guillaume, merci pour ce grain de folie qui te donne cet humour qui ne peut que faire rire ; Jérémy, merci pour ta bonne humeur et ton calme, ce soutien discret et toujours bienveillant ; Marion, ma compagne de pause thé, de séance de psychanalyse, de gym suédoise, et de tant d'autre chose, merci de m'avoir accordé autant de ton précieux temps, toi qui cours toujours entre Paris, Rennes et Tours, tu es sans doute la seule à avoir compris la vraie puissance de la Bretagne ; Mathieu, je te remercie entre autre pour tes découvertes culturelles, je t'associerais toujours à Alain Damasio. Je vous souhaite à tous bon courage pour vos fins de thèse !

Je remercie enfin ma famille, mes parents, mes grands-parents, mon frère et ma sœur. J'ai fini par réussir à la passer cette thèse ! Ce fut long, ce fut compliqué, mais j'ai réussi. Merci de m'avoir soutenue à chaque fois, de m'avoir questionnée sur mes choix, d'avoir répondu à mes questions et de m'avoir poussée à me poser les bonnes questions quand il fallait, mais de m'avoir toujours laissé faire ce qui était le mieux pour moi. Je n'aurais sans doute jamais pensé arrivé là où j'en suis actuellement sans vous tous.

Et je vais terminer ces remerciements en remerciant aussi toutes les boules de poils qui sans vraiment le savoir m'ont permis de finir cette thèse : ma petite puce Naïade, Vanille et Enora, Titinne et Réglisse.

# Table des matières

---

Résumé en français .....	2
Résumé en anglais .....	3
Remerciements .....	4
Table des matières .....	7
Table des figures.....	9
Tables des tableaux.....	11
Tables des annexes .....	13
Valorisation scientifique.....	14
Liste des abréviations .....	15
Introduction .....	18
Chapitre 1   État de l’art .....	26
1.1 Introduction à l’épidémiologie génétique .....	26
1.1.1 Marqueurs génétiques.....	26
1.1.2 Cartographie du génome.....	27
1.1.3 Identifier de la présence d’agrégation familiale .....	28
1.1.4 Analyses de ségrégation .....	30
1.1.5 Analyses de liaison génétique.....	31
1.2 Etudes d’association génétique .....	35
1.2.1 L’équilibre d’Hardy Weinberg .....	35
1.2.2 Le déséquilibre de liaison .....	36
1.2.3 Le principe des études d’association génétiques .....	38
1.3 HbA1c, mesure de l’équilibre glycémique .....	48
1.3.1 Le taux d’HbA1c : reflet de l’équilibre glycémique sur les trois derniers mois ..	48
1.3.2 Taux d’HbA1c et diabète.....	49
1.3.3 Facteurs non-génétiques influençant le taux d’HbA1c chez les personnes non- diabétiques.....	51
1.3.4 Les facteurs génétiques du taux d’HbA1c chez les personnes non-diabétiques...	53
Chapitre 2   Étude des propriétés statistiques du « General Regression Model » (GRM) dans le cas d’un trait quantitatif.....	65
2.1 Résumé.....	65
2.2 Article 1 (Publié en ligne en décembre 2019) .....	70

Chapitre 3   Facteurs influençant le taux d'HbA1c chez le patient souffrant de DT1 – Revue de la littérature	81
3.1 Résumé.....	81
3.2 Article 2 (publié en ligne en septembre 2018).....	87
Chapitre 4   Recherche de facteurs influençant le taux d'HbA1c dans une cohorte d'enfants français atteints de DT1 – Etude préliminaire .....	100
4.1 Introduction.....	100
4.2 Matériel et méthodes.....	102
4.2.1 Recrutement des patients et recueil des données .....	102
4.2.2 Transformation des données .....	104
4.2.3 Matériels biologiques et génotypage .....	107
4.2.4 Analyses statistiques.....	108
4.3 Résultats.....	112
4.3.1 Résultats des données cliniques à 2 ans de diabète .....	112
4.3.2 Analyse des données génétiques.....	121
4.3.3 Analyses des profils d'évolution d'équilibre glycémique .....	126
4.4 Discussion.....	133
Chapitre 5   Discussion générale .....	138
5.1 Le GRM.....	138
5.2 Le taux d'HbA1c chez le jeune patient souffrant de DT1 .....	143
Bibliographie .....	149
Annexes .....	175



# Table des figures

---

Figure 1 : Estimation du taux d'incidence (nombre de nouveaux cas par an pour 100 000) du DT1 chez les enfants de moins de 15 ans par la Fédération internationale du diabète (FID) (FID, 2019) .....	19
Figure 2 : Pathogenèse du DT1, d'après Katsarou <i>et al</i> 2017 (Katsarou <i>et al.</i> , 2017) .....	20
Figure 3 : Production de l'insuline et formation du peptide C, d'après le Beta Cell Biologie Consortium ( <a href="http://www.betacell.org/">http://www.betacell.org/</a> ) .....	22
Figure 4 : Odd-ratio des régions associées au diabète de type 1 (données extraites de GWAS Catalog le 10/02/2020) .....	23
Figure 5 : Crossing-over et gamètes .....	31
Figure 6 : Représentation d'un phénotype quantitatif pour les quatre principaux modes de transmission .....	33
Figure 7 : Proportion d'allèles partagés pour les différents statuts IBD .....	34
Figure 8 : Variation de la glycémie et du taux d'insuline durant la journée, en fonction du type de repas adapté de Daly <i>et al</i> 1998 (Daly <i>et al.</i> , 1998) .....	48
Figure 9 : Structure de l'HbA1c – tiré de Sherwani (Sherwani <i>et al.</i> , 2016) .....	49
Figure 10 : Risque de développer une complication en fonction du taux d'HbA1c .....	50
Figure 11 : Variation du taux d'HbA1c en fonction de l'âge et du genre .....	52
Figure 12 : Distribution du taux d'HbA1c en fonction de l'ethnie .....	52
Figure 13 : Différence du taux d'HbA1c entre individus non-diabétiques d'origines différentes vivant aux Etats-Unis, d'après Meigs <i>et al</i> 2014 (Meigs <i>et al.</i> , 2014) .....	53
Figure 14 : Marqueurs associés au taux d'HbA1c dans les différentes études disponibles sur GWAS Catalog .....	55
Figure 15 : Puissance du GRM et des tests MAX (Loley <i>et al.</i> , 2013) et allélique (Allelic) (Lee <i>et al.</i> , 2013) sur 1000 simulations de 1000 individus pour un seuil de 5E-8 en fonction de la MAF et de la PEV .....	67
Figure 16 : Flow-chart de l'étude GENDIR .....	103
Figure 17 : Variation du Taux d'HbA1c (A), de l'IDAA1c (B) et du taux d'HbA1c préajusté pour la DIR (C) à 2 ans et la loi normale associée .....	116
Figure 18 : Impact de la DIR sur le taux d'HbA1c à deux ans de diabète .....	117
Figure 19 : Impact de l'origine des parents sur le taux d'HbA1c à deux ans de diabète .....	117
Figure 20 : Analyse en composante principale, avec 25 individus pour lesquels nous ne possédons pas la donnée (noté Données Manquantes) .....	122
Figure 21 : Valeur du taux d'HbA1c au diagnostic, à 2, 5 et 10 ans dans notre cohorte .....	128
Figure 22 : Nuage de points du taux d'HbA1c aux différents temps d'analyse et le coefficient de corrélation de Pearson correspondant .....	129

Figure 23 : Nuage de points de la DIR aux différents temps d'analyse et le coefficient de corrélation de Pearson correspondant.....	129
Figure 24 : Évolution du nombre de patients en fonction de la catégorie du taux d'HbA1c ...	131
Figure 25 : Évolution du taux d'HbA1c des 132 patients qui possèdent des données à 2, 5 et 10 ans.....	132
Figure 26 : Modèle simplifié des différents facteurs impliqués dans la régulation du taux d'HbA1c chez les patients souffrant de DT1.....	143

# Tables des tableaux

---

Tableau 1 : Codage du génotype en fonction du modèle génétique considéré.....	41
Tableau 2 : Codage des variables pour les tests génotypiques disponibles sur PLINK (Chang <i>et al.</i> , 2015; Purcell <i>et al.</i> , 2007) .....	42
Tableau 3 : Codage des génotypes .....	44
Tableau 4 : Contrastes des paramètres $\beta$ .....	45
Tableau 5 : Estimation de l'héritabilité du taux d'HbA1c chez des patients non-diabétiques ...	54
Tableau 6 : GWAS et méta-analyses étudiant le taux d'HbA1c disponibles sur GWAS Catalog .....	56
Tableau 7 : Régions considérées comme étant glycémiques et qui ont été identifiées dans au moins deux études, ou avec un coefficient de régression supérieur à 0.05 .....	60
Tableau 8 : Régions considérées comme étant érythrocytaires et qui ont été identifiées dans au moins deux études, ou avec un coefficient de régression supérieur à 0.05 .....	62
Tableau 9 : Régions considérées comme étant autre qui ont été identifiées dans au moins deux études, ou avec un coefficient de régression supérieur à 0.05 .....	64
Tableau 10 : Codage des variables en fonction du génotype .....	66
Tableau 11 : Facteurs identifiés comme étant associés à un moins bon équilibre glycémique (Neylon <i>et al.</i> , 2013).....	82
Tableau 12 : Facteurs démographiques à deux ans de diabète dans la cohorte GENDIR.....	113
Tableau 13 : Facteurs liés au diagnostic du DT1 dans la cohorte GENDIR .....	114
Tableau 14 : Auto-anticorps (AAC) mesurés au moment du diagnostic de DT1.....	114
Tableau 15: Facteurs liés au traitement du DT1 à deux ans de diabète dans la cohorte GENDIR .....	115
Tableau 16 : Autres facteurs collectés dans la cohorte GENDIR.....	115
Tableau 17: Données glycémiques à 2 ans de DT1 .....	116
Tableau 18 : Seuil de significativité des analyses univariées de l'IDAA1C, du taux d'HbA1c ajusté pour la DIR et du taux d'HbA1c préajusté pour la DIR à 2 ans de DT1 .....	118
Tableau 19 : Résultats des régressions multivariées et pas à pas descendantes .....	120
Tableau 20 : Descriptif de la cohorte GENIR à 2 ans et du sous-échantillon génotypé .....	123
Tableau 21 : Marqueurs identifiés comme étant associés au taux d'HbA1c dans les différentes publications, publiées avant septembre 2019 .....	124
Tableau 22 : Analyse des marqueurs associés au taux d'HbA1c préajusté pour la DIR chez les patients DT1 dans notre échantillon .....	126
Tableau 23 : Évolution des différentes variables au cours du suivi .....	127

Tableau 24 : Développement des complications et des comorbidités après 10 ans de maladie..... 130  
..... 130

# Tables des annexes

---

Annexe 1 : Données supplémentaires de l'étude des propriétés statistiques du GRM (Gloaguen <i>et al.</i> , 2019).....	175
Annexe 2 : Données supplémentaire de la revue systématique des facteurs influençant le taux d'HbA1c chez le jeune patients DT1 (Gloaguen <i>et al.</i> , 2018) .....	188
Annexe 3 : Consentements envoyés aux patients et/ou à leur parents en fonction de leur âge au moment de l'étude.....	203
Annexe 4 : Cahier de recueil de données (CRF) de l'étude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une collection de patients atteints de diabète de type I (GENDIR) .....	217
Annexe 5 : Phénotype et processus biologiques associés aux gènes testés dans l'étude GENDIR .....	233

# Valorisation scientifique

---

## ARTICLES :

- Gloaguen E, Bendelac N, Nicolino M, Julier C, Mathieu F. A systematic review of non-genetic predictors and genetic factors of glycated haemoglobin in type 1 diabetes one year after diagnosis. *Diabetes Metab Res Rev.* 2018;34(8):e3051. doi:10.1002/dmrr.3051
- Gloaguen E, Dizier MH, Boissel M, et al. General regression model: A "model-free" association test for quantitative traits allowing to test for the underlying genetic model [published online ahead of print, 2019 Dec 13]. *Ann Hum Genet.* 2019;10.1111/ahg.12372.

## POSTERS :

- Emilie Gloaguen, Dizier MH, Julier C, Mathieu F, General Regression Model: a powerful "model free" association test for quantitative traits allowing to test for the underlying genetic model (European Mathematical Genetics Meeting 2015, Brest).
- Emilie Gloaguen, M-H. Dizier, N. Bendelac, C. Julier, M. Nicolino, F. Mathieu, Factors influencing glycated hemoglobin in type 1 diabetes patients (International Symposium of Insulin Receptor and Insulin Action 2017).

# Liste des abréviations

---

AAC : Auto-Anticorps

AAC anti-GAD : Auto-Anticorps anti Glutamate décarboxylase ou *glutamic acid decarboxylase* en anglais

ADA : Association Américaine des Diabètes ou *American Diabetes Association*

ADN : Acide Désoxyribonucléique

ANOVA : Analyse de la variance pour *ANalyse Of VAriance*

BCR : Récepteur des lymphocyte B

BNC2 : *Basonucline 2*

CBC-Biotec : Centre de Biotechnologie Cellulaire et Biothèque

CCR5 : *C-C chemokine receptor type 5*

CCRL2 : *CC chemokine receptor-like 2*

CGM : continuous glucose monitoring

cM : centiMorgan

CMH : Complexe Majeur d'Histocompatibilité

CNV : *Copy Number Variant* en anglais

CSII : pour l'anglais *Continuous Subcutaneous Insulin Infusion*, correspondant à un traitement sous pompe

CSP : catégorie socio-professionnel

CTLA4 : cytotoxic T-lymphocyte-associated protein 4

DC : Cellules dendritiques

DCCT : Diabetes Control and Complications Trial

D.E.S.I.R. : *Data from an Epidemiological Study on the Insulin Resistance Syndrome* ou Données Épidémiologiques sur le Syndrome d'InsulinoRésistance

DIAGRAM : *DIAbetes Genetics Replication And Meta-analysis*

DIR : Dose d'insuline requise par jour ou *Daily Insulin Requirement*

DKA : Acidocétose diabétique ou *diabetic ketoacidosis* en anglais

DPV : *Diabetes-Patienten-Verlaufsdokumentation* en allemand

DSI : département système d'information

DT1 : Diabète de type 1

DT2 : Diabète de type 2

DZ : jumeaux dizygotes

eCRF : Cahier de Recueil électronique de données, pour l'anglais *electronic Case Report Form*

EDIC : *Epidemiology of Diabetes Interventions and Complications*

ERBB3 : erb-b2 receptor tyrosine kinase 3

ExAC : *Exome Aggregation Consortium*

FDR : *False Discovery Rate*

FID : Fédération internationale du diabète

FinnDiane : *Finnish Diabetic Nephropathy*

FPG : *fasting plasma glucose* (glycémie à jeun)

GENDIR : Étude des facteurs GENétiques impliqués dans le contrôle de la Dose d'Insuline journalière Requête pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type 1

GRM : *General Régression Model* ou en français modèle général de régression

GSC : *Goosecoid Protein*

GWAS : *Genome-Wide Association Study* ou étude d'association pangénomique

H0 : hypothèse nulle

HapMap : Haplotype Map

HbA1c : Hémoglobine A1c ou glyqué

HW : Équilibre d'Hardy Weinberg ou *Hardy Weinberg Equilibrium* en anglais

IBD : *Identical By Descent* en anglais

IBS : *Identical By State* en anglais

ICA : Autoanticorps anti-ilots

IDAA1C : *Insulin Dose-Adjusted A1C* en anglais

IFCC : *International Federation of Clinical Chemistry and Laboratory Medicine*

IL6 : Interleukine 6

IMC : Indice de Masse Corporelle (*BMI* pour *Body Mass Index* en anglais)

IMC-SDS : IMC standardisé

Insee : Institut national de la statistique et des études économiques

ISPAD : *International Society for Pediatric and Adolescent Diabetes*

Lasso : *Least Absolute Shrinkage and Selection Operator* en anglais

LD : déséquilibre de liaison pour l'anglais *linkage disequilibrium*

LOD-score : *logarithm of odds*, logarithme des probabilités en anglais

LR : *Linear Régression* ou régression linéaire

LR\_ADD : LR sous un modèle additif

LR\_DOM : LR sous un modèle dominant

LR\_GENO : LR sous un modèle génotypique

LR\_HETHOM : LR avec un codage hétérozygote/homozygote

LR\_REC : LR sous un modèle récessif

LRP5 : *LDL receptor-related protein 5*

MAF : Fréquence d'Allèle Mineur ou en anglais *Minor Allele Frequencies*

MAX : Maximum of the standardized optimal tests

MODY : *Maturity Onset of the Young*

MZ : jumeaux monozygotes

NGS : *Next-Generation Sequencing*

NHANES : *National Health and Nutrition Examination Survey*

NHGRI : *National Human Genome Research Institute* (<https://www.genome.gov/>)

NPAI : n'habitant plus à l'adresse indiquée

NPH : *Neutral Protamine Hagedorn*

OMS : Organisation mondiale de la santé

OR : *odd-ratio* en anglais ou rapport des cotes

PCA : Analyse en Composante Principales ou *Principal Components analysis* en anglais

PEV : Pourcentage de variance expliqué ou en anglais *Percentage of Explained Variance*

PharmGKB : *Pharmacogenomics Knowledgebase*

PTPN22 : *Protein Tyrosine Thosphatase, Ton receptor type 22*

RASS : *Renin–Angiotensin System Study*



RedCap : pour l'anglais *Research Electronic Data Capture*  
RFLP : restriction fragment length polymorphism en anglais  
RXFP2 : *Relaxin Family Peptide Receptor 2*  
SES : Statut Socio-économique  
SLC30A8 : *Solute Carrier Family 30 Member 8*  
SNIIRAM : système national d'information inter-régimes de l'Assurance Maladie  
SNP : Single Nucleotide Polymorphism en anglais  
SORCS1 : *sortilin-related vacuolar protein sorting 10 domain containing receptor 1*  
STR : *Short Tandem Repeat* en anglais  
TCR : Récepteur des lymphocytes T  
TLR3 : *Toll-like receptor 3*  
TNF : *tumor necrosis factor*  
VNTR : *Variable Number Tandem Repeat* en anglais  
WDR72 : *WD repeat-containing protein 72*

# Introduction

---

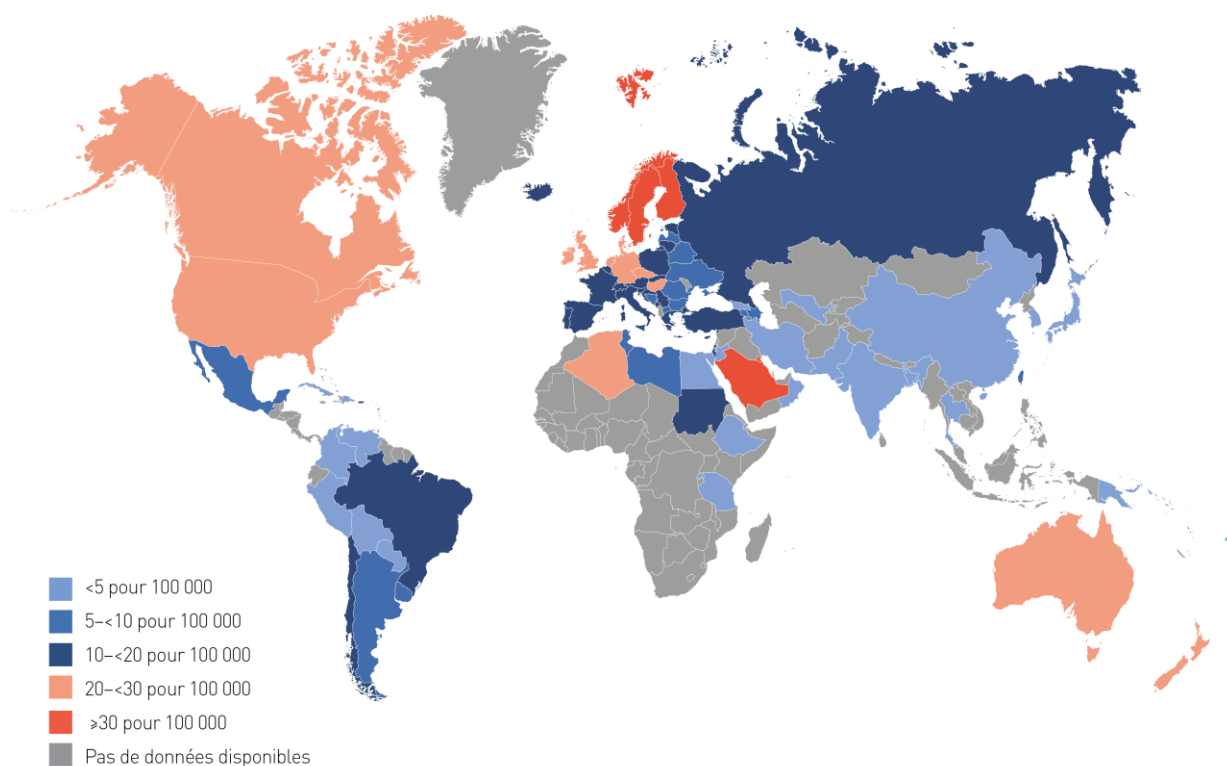
Historiquement, le diabète est défini comme une maladie caractérisée par une augmentation de la production d'urine, et regroupe deux types de maladies : les diabètes sucrés, caractérisés par la présence de glucose dans les urines et qui sont généralement dus à un défaut de reconnaissance ou de sécrétion de l'insuline et les diabètes insipides, sans présence de glucose dans les urines et qui sont dus à un défaut de sécrétion ou de reconnaissance de l'hormone antidiurétique. Aujourd'hui, l'Organisation Mondiale de la Santé (OMS) définit le mot diabète uniquement à la forme dite sucrée, beaucoup plus fréquente (OMS, 2016).

L'Association Américaine des Diabètes (ADA) définit quatre formes de diabète (ADA, 2019b) :

- Le diabète de type 1 (DT1) (autrefois appelé diabète insulino-dépendant ou juvénile) qui est dû à la destruction auto-immune des cellules  $\beta$  des îlots de Langerhans, cellules qui produisent et sécrètent l'insuline ;
- Le diabète de type 2 (DT2) (autrefois dit diabète non insulino-dépendant ou diabète de la maturité) qui résulte d'une diminution progressive de la sécrétion d'insuline et/ou d'une résistance à l'insuline. C'est la forme de diabète la plus fréquente ;
- Le diabète gestationnel qui apparaît uniquement durant la grossesse de façon transitoire ;
- Les autres formes de diabète, associés à des caractéristiques particulières, dont les formes les plus courantes sont syndromiques, c'est-à-dire associées à d'autres symptômes, ainsi que les diabètes MODY (pour *Maturity Onset of the Young*).

De 5 à 10% des personnes souffrant de diabète dans le monde sont atteints de DT1 (ADA, 2019b). L'estimation de l'incidence du DT1 pour les moins de 15 ans dans le monde en 2019 est représentée dans la Figure 1. En France, cette incidence a été estimée à 13 à 14 cas pour 100 000 en 2004, avec une augmentation de 3% entre 1988 et 2004 (Barat, 2016). Ce taux annuel a continué d'augmenter entre 2010 et 2015, en passant de 16 cas en 2010 à 19 cas en 2015 pour 100 000 (Piffaretti *et al.*, 2019). Cette évolution est aussi visible dans l'ensemble de l'Europe, entre 1989 et 2008, avec une augmentation de la fréquence moyenne de DT1 de 3 à 4% par an (Patterson *et al.*, 2012). Aux Etats-Unis, l'incidence du DT1 a également augmentée

entre 2001 et 2015 chez les enfants âgés entre 0 et 19 ans. En revanche, cette incidence a eu tendance à diminuer pour les adultes âgés entre 20 et 64 ans (Rogers, Kim, Banerjee, & Lee, 2017).

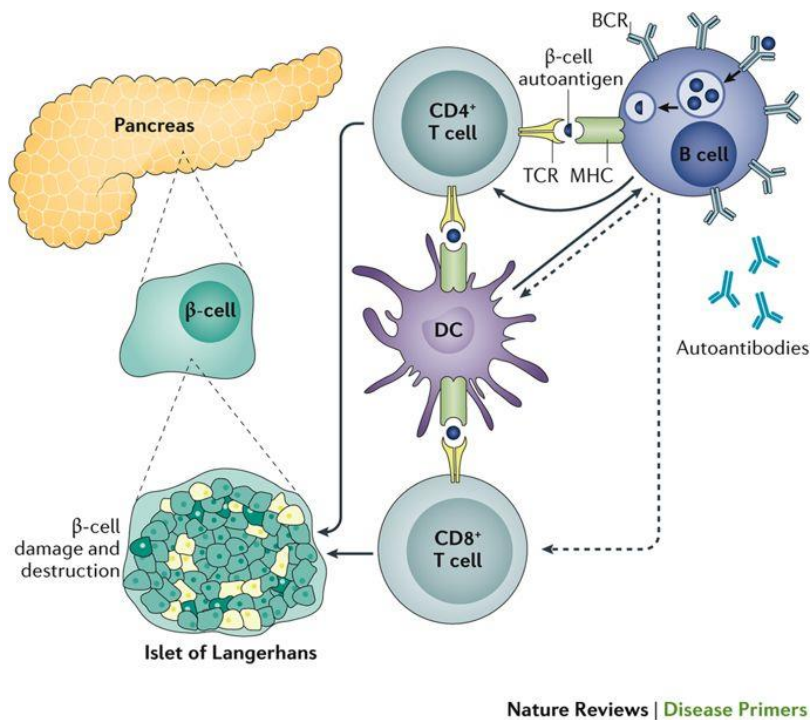


**Figure 1 :** Estimation du taux d'incidence (nombre de nouveaux cas par an pour 100 000) du DT1 chez les enfants de moins de 15 ans par la Fédération internationale du diabète (FID) (FID, 2019)

En 2017, Katsarou *et al* ont publié une revue sur le DT1, reprenant notamment les mécanismes conduisant à la destruction des cellules  $\beta$  (Katsarou *et al.*, 2017). Ces mécanismes sont décrits dans la Figure 2, tirée de cette revue. L'origine de la production des premiers auto-anticorps (AAC) reste encore aujourd'hui inconnue, mais l'hypothèse de l'intervention conjointe d'infection virale et de facteurs environnementaux dans un contexte génétique particulier semble la plus probable (Szablewski, 2014).

Les symptômes classiques du DT1 comprennent généralement, en plus d'une hyperglycémie, une polyurie et une polydipsie, c'est-à-dire une augmentation de la consommation d'eau et de production d'urine ainsi qu'une perte de poids associée à une fatigue importante (ADA, 2019b). Dans environ un tiers des cas, le patient présente aussi une acidocétose diabétique (DKA pour *diabetic ketoacidosis* en anglais) qui est due à la production de corps cétoniques suite à

l'utilisation de graisse pour parer à l'absence de glucose (Dabelea *et al.*, 2014). Dans les cas les plus graves, la DKA peut entraîner un coma.



**Figure 2 :** Pathogénèse du DT1, d'après Katsarou *et al* 2017 (Katsarou *et al.*, 2017)

Les lymphocytes B (B cell) activés interagissent avec les lymphocytes T CD4+ (CD4+ T cell) et CD8+ (CD8+ cell) ainsi qu'avec les cellules dendritiques (DC). La présentation de l'antigène par les lymphocytes B et les DC entraîne l'activation des lymphocytes T spécifiques des cellules  $\beta$ . De plus, l'exposition des lymphocytes B à des auto-antigènes de cellule  $\beta$  conduit à la production d'auto-anticorps ciblant les îlots de Langerhans, qui servent de biomarqueurs à la maladie. Les flèches en pointillés indiquent les interactions potentielles entre les lymphocytes B et les lymphocytes T CD8+ et entre les lymphocytes B et les cellules dendritiques (DC). BCR, récepteur des cellules B ; TCR, récepteur des cellules T.

La confirmation du diagnostic de diabète se fait généralement par la mesure de la glycémie à jeun et du taux de glucose après un test de tolérance au glucose. La confirmation de l'état auto-immun de la maladie se fait par la détection d'un ou de plusieurs auto-anticorps (AAC) dans le sang au moment du diagnostic (Szablewski, 2014) :

- Les AAC anti-îlots (ICA) : présents dans la majorité des cas ;
- Les AAC anti-GAD (Glutamate décarboxylase ou *glutamic acid decarboxylase* en anglais) : dirigés contre une enzyme ubiquitaire, exprimée au niveau pancréatique et présents jusqu'à 10 ans avant le diagnostic ;
- Les AAC anti-IA2 : dirigés contre une protéine membranaire présente dans les cellules  $\beta$  ;
- Les AAC anti-insuline : directement dirigés contre l'insuline ;

- Les AAC anti-ZNT8 : dirigés contre un transporteur de zinc qui est présent à la surface des cellules  $\beta$ , codé par le gène SLC30A8 (*Solute Carrier Family 30 (Zinc Transporter), Member 8*).

Tous les patients n'expriment pas l'ensemble de ces AACs. La présence d'un seul AAC permet de classer les patients comme étant atteints de DT1 (Kawasaki, 2014).

Le patient souffrant de DT1 doit recevoir des injections d'insuline exogène afin de réguler sa glycémie et d'éviter les crises d'hyperglycémie qui sont à l'origine de complications à long terme, telles que les neuropathies ou les néphropathies (Katsarou *et al.*, 2017). Les premiers types de traitement consistaient à injecter une ou deux doses d'insuline dans la journée. Les pompes à insuline sous-cutanées permettent une délivrance d'insuline en continue afin de reproduire au mieux la fonction normale du pancréas (Alsaleh, Smith, Keady, & Taylor, 2010). Dans les années 90, des études telles que le DCCT (*Diabetes Control and Complications Trial Research Group*) ont permis de montrer qu'un traitement plus intensif consistant à plus de 2 injections par jour ou à l'utilisation d'une pompe à insuline permettait d'améliorer l'équilibre glycémique et de prévenir les complications (DCCT, 1994). Récemment, une revue a repris les points clefs de l'histoire et du mécanisme de la pompe à insuline (ou *Continuous subcutaneous insulin infusion* (CSII) en anglais) (Mavinkurve, Quinn, & O'Gorman, 2016). L'utilisation des pompes semblent être en constante augmentation, passant de seulement 4% des patients pédiatriques en 2006 à plus de 42% en 2012. Mais aujourd'hui, le schéma le plus classique reste un modèle appelé « basal-bolus » qui consiste en une ou deux doses d'insuline lente ou intermédiaire au moment du coucher et du réveil et en une dose rapide à chaque repas en fonction de la glycémie mesurée avant le repas. Ce système permet d'avoir une certaine souplesse quant à l'heure des repas ou à l'activité physique et consiste à faire des injections à l'aide d'un stylo à insuline par type d'insuline.

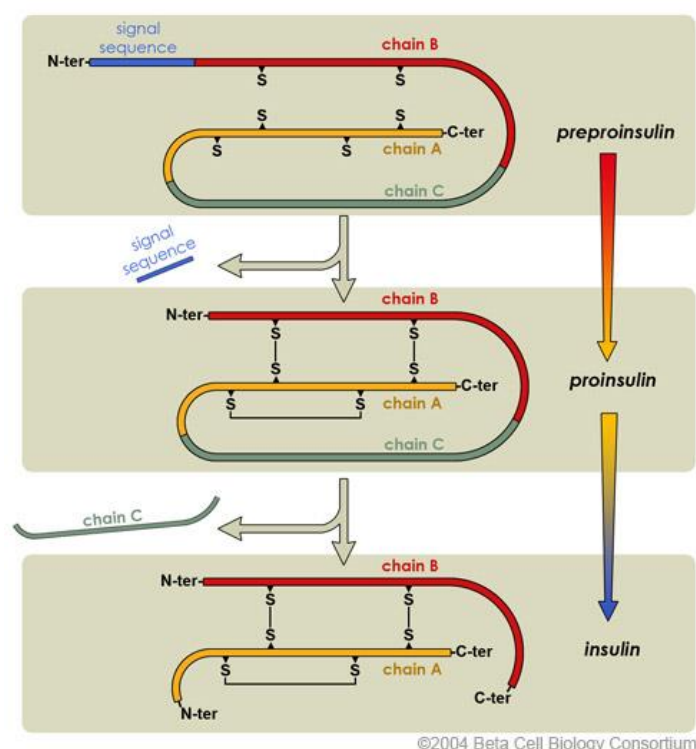
Il existe plusieurs types d'insuline utilisée pour le traitement du diabète, référencées dans une revue récente (Mathieu, Gillard, & Benhalima, 2017), dont la forme, la durée d'activité et d'absorption sont différentes, :

- Des insulines dites à action rapide (aspart) : action rapide, au bout d'une trentaine de minutes avec un pic au bout d'une heure, et d'une durée d'action d'environ 6 heures ;

- Des insulines intermédiaires (NPH ou *Neutral Protamine Hagedorn*) : action pendant 12 à 16 heures ;
- Des insulines à action lente (Glargine, Detemir) : action pouvant aller jusqu'à 24 heures.

Il est aussi possible d'effectuer des mélanges d'insuline qui regroupent des insulines rapides et des insulines lentes dans un même stylo afin de n'effectuer qu'une seule injection. Le choix du type d'insuline va dépendre des caractéristiques du patient au niveau de son profil glycémique, mais aussi de son âge ou de son mode de vie.

Au début du traitement, la majorité des patients présente une rémission partielle, qui peut durer de quelques mois à un peu plus de un an et qui se caractérise par un faible besoin en insuline (Abdul-Rasoul, Habib, & Al-Khouly, 2006). Ceci est dû à l'activité résiduelle des cellules- $\beta$  des îlots de Langerhans qui n'ont pas encore été détruites par le système immunitaire. Un des biomarqueurs permettant de détecter la période de rémission partielle est le peptide-C, qui permet d'avoir une idée de la production d'insuline endogène, (VanBuecken & Greenbaum,



**Figure 3 :** Production de l'insuline et formation du peptide C, d'après le Beta Cell Biologie Consortium

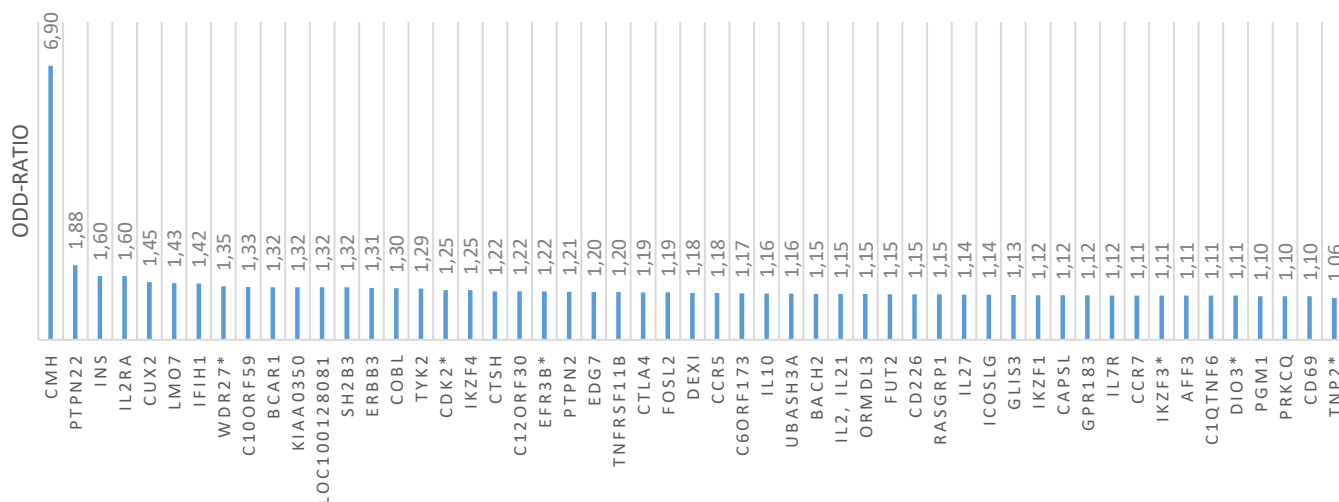
(<http://www.betacell.org/>)

La pré-pro-insuline, composée de 3 chaînes A, B et C et d'une séquence de signal ; est synthétisée par les cellules  $\beta$  des îlots de Langerhans, une première étape de maturation, consistant à retirer la séquence signal, va permettre de former la pro-insuline qui sera ensuite clivée en insuline, composée des chaînes A et B, et en peptide C.

2014) puisqu'il correspond au peptide de liaison entre les deux chaînes actives de l'insuline présentes dans la pro-insuline (Figure 3). Cette période est particulièrement étudiée car elle pourrait permettre la mise en place de traitements permettant de limiter la destruction des cellules- $\beta$  (Fonolleda, Murillo, Vazquez, Bel, & Vives-Pi, 2017).

Le DT1 est une maladie complexe dont les causes sont à la fois génétiques (Pociot & Lernmark, 2016), épigénétiques (Wang & Lu, 2017; Xiang, Yang, Chang, & Lu, 2017) et environnementales (Rewers & Ludvigsson, 2016). Parmi les facteurs environnementaux identifiés, les principaux facteurs candidats sont les infections, le régime alimentaire, les toxines et le stress *in utero*, pendant la période périnatale ou pendant la petite enfance (Rewers & Ludvigsson, 2016).

Des analyses de gènes candidats ainsi que des études pangénomique ou GWAS (pour *Genome-Wide Association Study*) ont permis de mettre en évidence environ soixante *loci* associés au DT1. Les analyses d'association ont été effectuées à l'aide de régression logistique, permettant ainsi d'estimer des *odd-ratios* (OR), qui mesurent l'effet du marqueur sur le DT1. La base de données GWAS Catalog permet de récupérer l'ensemble des OR calculés pour les marqueurs significativement associés au DT1 ( $p < 5E-8$ ), ces OR sont reportés dans la Figure 4. La plupart de ces *loci* ont des effets faibles (Pociot & Lernmark, 2016), à l'exception de la région du complexe majeur d'histocompatibilité (CMH), le *locus* majeur du DT1 (WTCCC, 2007).



**Figure 4 :** Odd-ratio des régions associées au diabète de type 1 (données extraites de GWAS Catalog le 10/02/2020)

Les régions, dont le gène reporté a une étoile, correspondent à des régions pour lesquelles plusieurs gènes ont été reportés dans les articles montrant leur associations (Barrett *et al.*, 2009; Bradfield *et al.*, 2011; Cooper *et al.*, 2008; Grant *et al.*, 2009; Hakonarson *et al.*, 2007; Hakonarson *et al.*, 2008; Huang, Ellinghaus, Franke, Howie, & Li, 2012; Onengut-Gumuscu *et al.*, 2015; Todd *et al.*, 2007; Wallace *et al.*, 2010; 2007).

L'ensemble de ces facteurs génétiques n'explique qu'une fraction de l'héritabilité du DT1. La fraction non expliquée, dite « héritabilité manquante », peut être due aux limites des méthodologies statistiques utilisées pour les GWAS (Genin & Clerget-Darpoux, 2015).

Malgré les avancées technologiques au cours des 20 dernières années, le contrôle de la glycémie est toujours un défi de taille pour les enfants atteints de DT1 et leurs familles. Certains patients n'atteignent jamais un équilibre glycémique correct, ce qui entraîne des complications sévères à long terme (néphropathie, neuropathie, rétinopathie, etc.). L'équilibre glycémique est estimé par le taux d'hémoglobine glyquée, noté HbA1c.

Le taux d'HbA1c est un phénotype quantitatif complexe. Sa variabilité peut être expliquée par différents facteurs (génétiques ou non) et par les interactions entre ces facteurs. Cette thèse a pour objectif principal d'identifier les facteurs génétiques influençant l'équilibre glycémique en prenant en compte les facteurs non-génétiques connus. Afin d'atteindre cet objectif, ce projet comporte trois volets, qui seront présentés dans les chapitres suivants.

Le Chapitre 1 est un état de l'art, présentant les différentes connaissances actuelles sur les méthodes d'analyse génétique et sur le taux d'HbA1c.

Le Chapitre 2 présente les propriétés statistiques du GRM (*General Regression Model* ou modèle général de régression). En effet, pour rechercher les facteurs génétiques impliqués dans le taux d'HbA1c, nous avons choisi d'utiliser le GRM dont nous avons étudié les propriétés statistiques dans le cadre de l'étude d'un trait quantitatif à l'aide d'une étude de simulation ainsi que d'une application sur les données d'HbA1c de la cohorte D.E.S.I.R. (*Data From an Epidemiological Study on the Insulin Resistance Syndrome*). Le GRM, déjà décrit comme plus puissant que les méthodes classiques dans le cas d'un trait binaire (Dizier, Demenais, & Mathieu, 2017), a été comparé à des méthodes de régression linéaire classiquement utilisées, ainsi qu'à différents tests proposés plus récemment. Ce travail a fait l'objet d'une publication dans *Annals of Human Genetics* (Gloaguen *et al.*, 2019), ainsi que d'un poster à l'*European Mathematical Genetics Meeting* en 2015.

Le Chapitre 3 présente une revue systématique de la littérature sur les facteurs influençant le taux d'HbA1c chez les enfants atteints de DT1. En effet, de nombreuses études des facteurs démographiques, socio-économiques et cliniques qui influencent le taux d'HbA1c ont été



menées depuis 2010, nécessitant d'en faire une synthèse afin d'identifier les principaux facteurs influençant le taux d'HbA1c. Ce travail a conduit à l'écriture d'une revue systématique de la littérature publiée entre 2011 et 2016 des facteurs influençant le taux d'HbA1c chez l'enfant DT1. Ce travail a fait l'objet d'une publication dans *Diabetes/Metabolism: Research and Reviews* (Gloaguen *et al.*, 2018) ainsi que d'un poster présenté au *International Symposium on Insulin Receptor and Insulin action* et à l'école d'été associée *DZD Diabetes Research School* en 2017.

Le Chapitre 4 présente l'étude préliminaire d'une cohorte rétrospective de jeunes patients français atteints de DT1 (Projet GENDIR pour Etude des facteurs GENétique impliqués dans le contrôle de la Dose d'Insuline journalière Requisite). Ce projet a pour but de rechercher les facteurs à risque, génétiques ou non, impliqués dans le contrôle de l'équilibre glycémique à 2 ans, et d'étudier son évolution à 5 et 10 ans. J'ai participé à la collecte des données en collaboration avec le service d'endocrinologie pédiatrique de l'Hôpital Femme-Mère-Enfant des Hospices Civils de Lyon (Dr Nathalie Bendelac et Pr M. Nicolino) et réalisé l'ensemble des analyses préliminaires.

Pour conclure, l'ensemble de ces travaux seront discutés et différentes perspectives seront proposées dans le Chapitre 5.

# Chapitre 1 | État de l'art

---

## 1.1 Introduction à l'épidémiologie génétique

Les maladies complexes, comme le DT1, ont une étiologie multifactorielle expliquée par l'effet conjoint de plusieurs gènes, de l'environnement et d'interactions de type gène-gène et/ou gène-environnement. La localisation des *loci* responsables de la susceptibilité génétique d'un phénotype ou d'une maladie a été rendue possible grâce à la découverte des marqueurs génétiques, qui sont des polymorphismes (variations entre individus) de l'ADN (acide désoxyribonucléique) dont les positions exactes sont connues dans le génome.

### 1.1.1 Marqueurs génétiques

Il existe différentes catégories de marqueurs génétiques, pouvant être bi-alléliques ou multi-alléliques, rares ou fréquents dans la population. La plupart des marqueurs génétiques sont des polymorphismes de séquence.

Historiquement, les premiers marqueurs génétiques étudiés ont été les polymorphismes de longueur des fragments de restriction, ou RFLP (pour *Restriction Fragment Length Polymorphism* en anglais) (Kan & Dozy, 1978). Il s'agit d'une variation d'un nucléotide<sup>1</sup> présent au niveau d'un site de coupure d'enzyme de restriction (marqueurs bi-alléliques). La présence (ou l'absence) de la variation détermine si l'enzyme coupera (ou non) l'ADN au niveau du site de restriction. Ces marqueurs sont peu nombreux dans le génome et nécessitent d'utiliser une grande quantité d'ADN pour le génotypage.

Dans les années 1990, les séquences répétées en tandem (VNTR pour *Variable Number Tandem Repeat* en anglais) ont beaucoup été utilisées. Ce sont des polymorphismes multi-alléliques,

---

<sup>1</sup> Unité élémentaire de l'ADN, composée d'une des quatre bases nucléiques : l'adénine (A), la thymine (T), la cytosine (C) et la guanine (G)

dont le nombre et la taille sont très variables d'un individu à l'autre. On distingue deux types de VNTRs, les plus utilisés étant les microsatellites ou STR (pour *Short Tandem Repeat* en anglais). Environ 5200 microsatellites ont été caractérisés dans le génome humain et sont encore utilisés aujourd'hui car ils sont très polymorphes d'une personne à l'autre.

Les variants structuraux correspondent à des délétions, insertions, duplications, inversions ou translocations<sup>2</sup> d'un segment d'ADN de taille supérieure ou égale à 1kb. Ils ont été très utilisés pour les études de cytogénétique<sup>3</sup>, puisqu'ils peuvent être observables à l'échelle microscopique ou sub-microscopique, dans le cadre de réalisation de caryotype ou à l'aide de méthode d'hybridation *in situ* par des sondes.

Plus récemment, les CNVs (pour *Copy Number Variant* en anglais) ont été identifiés. Ce sont des mini-délétions ou duplications de petits fragments du génome (entre 1kb et 3Mb). Les CNVs représentent entre 4.8 et 9.5% du génome humain (Zarrei, MacDonald, Merico, & Scherer, 2015) et sont associés à la notion de dosage génique (McCarroll & Altshuler, 2007).

L'évolution des techniques de biologie moléculaire depuis les années 2000 a vu l'essor du génotypage de polymorphismes d'un seul nucléotide appelés SNP (pour *Single Nucleotide Polymorphism* en anglais). Les SNPs sont des polymorphismes d'un seul nucléotide et sont généralement bi-alléliques. Les SNPs ont l'avantage de couvrir l'ensemble du génome à la fois dans les régions codantes et non-codantes.

### 1.1.2 Cartographie du génome

Le projet HapMap (pour *Haplotype Map*) a permis de caractériser un grand nombre de SNPs (localisation, allèles, fréquences, etc.) dans 11 populations humaines. Environ 4 millions de SNPs ont été génotypés durant les trois phases de ce projet (Octobre 2005 (HapMap, 2005), Octobre 2007 (Frazer *et al.*, 2007) et Septembre 2010 (Altshuler *et al.*, 2010)). Une carte du génome humain a ainsi été développée et mise à disposition des chercheurs du monde entier (HapMap,

---

<sup>2</sup> Échange réciproque de matériel chromosomique entre des chromosomes non homologues

<sup>3</sup> Etude de la structure et du nombre de chromosomes

2003). L'ensemble de ces informations (fréquences alléliques, séquences, etc.) des SNPs est disponible pour la communauté scientifique dans des bases de données accessibles en ligne, dont dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>) (Sherry *et al.*, 2001). Aujourd'hui, les avancées technologiques ont permis d'automatiser le génotypage à l'aide de « puces », permettant de génotyper un grand nombre de sujets pour un grand nombre de SNPs (jusqu'à 2.5 millions) et pour un coût raisonnable.

Les techniques de séquençage du génome permettent de connaître l'ensemble de la séquence d'une région ou de régions données ou de l'ensemble du génome. Le premier séquençage d'un génome humain complet a été terminé en 2003 et a coûté environ 3 millions de dollars. Aujourd'hui, le séquençage est automatisé et coûte environ 1000 dollars pour un génome.

En 2008, le projet *1000 Genomes* a été mis en place pour identifier et cartographier plus finement l'ensemble des variations génétiques humaines (Abecasis *et al.*, 2010). Ce projet portant sur un grand nombre d'individus (supérieur à 1000) de différentes populations humaines a permis d'identifier 88 millions de polymorphismes dans le génome, dont 84,7 millions de SNPs (Auton *et al.*, 2015). Des études plus récentes, telles que le *Exome Aggregation Consortium* (ExAC) (Karczewski *et al.*, 2017; Lek *et al.*, 2016), ont permis de mettre en évidence de nouveaux polymorphismes (fréquents ou rares) dans différentes populations.

Pour identifier et caractériser la composante génétique de la variabilité d'un trait ou d'une maladie, il faut tout d'abord démontrer qu'il existe une agrégation familiale due à une composante génétique, puis localiser le ou les *loci* impliqué(s) en utilisant les marqueurs génétiques et enfin préciser le (ou les) effets de cette composante génétique.

### 1.1.3 Identifier de la présence d'agrégation familiale

#### 1.1.3.1 Le risque de récurrence

L'importance de l'agrégation familiale est souvent estimée par le risque relatif de récurrence. Pour un apparenté donné, ce risque correspond au rapport entre la prévalence de la maladie chez ses apparentés ( $K_S$ ) par rapport à la prévalence de la maladie dans la population générale ( $f_M$ ). Le risque relatif de récurrence entre germains (frères et sœurs) est le plus utilisé, noté  $\lambda_s$ , s pour *sib* en anglais (Risch, 1990) :

$$\lambda_S = \frac{K_S}{f_M}$$

Un risque relatif supérieur à 1 correspond à la présence d'agrégation familiale. Plus ce risque est élevé, plus l'agrégation familiale est forte. Cependant, cette agrégation n'est pas forcément due à un facteur génétique mais peut aussi être due à un facteur d'environnement familial partagé.

### 1.1.3.2 L'héritabilité

Le calcul de l'héritabilité a été proposé pour quantifier la part de la variance du phénotype due à l'effet des facteurs génétiques et environnementaux sur le phénotype étudié. La variance du phénotype ( $\sigma_P^2$ ) est décomposée en variance génotypique ( $\sigma_G^2$ ) et en variance environnementale ( $\sigma_E^2$ ), avec  $\sigma_P^2 = \sigma_G^2 + \sigma_E^2$ . La variance génétique peut elle-même être décomposée en trois termes : une variance additive ( $\sigma_A^2$ ), une variance de dominance ( $\sigma_D^2$ ) et une variance d'interaction ( $\sigma_I^2$ ), avec  $\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2$ .

L'héritabilité au sens strict ( $h^2$ ) est estimée par le rapport entre la variance additive et la variance phénotypique :  $h^2 = \frac{\sigma_A^2}{\sigma_P^2}$ . L'héritabilité au sens large ( $H^2$ ) est estimée par le rapport entre la variance génétique et la variance phénotypique :  $H^2 = \frac{\sigma_G^2}{\sigma_P^2}$ . Il n'existe aucun consensus sur les méthodes d'estimation de l'héritabilité, de  $\sigma_G^2$ , de  $\sigma_A^2$  et de  $\sigma_P^2$ . Dans la plupart des études,  $H^2$  est estimée sans prendre en compte les effets d'interaction, ni les effets d'environnement.

L'étude de cohortes de jumeaux, monozygotes (MZ) et dizygotes (DZ) a été proposée pour dissocier la part de variabilité due aux facteurs génétiques et celle due aux facteurs environnementaux. Pour un phénotype binaire, l'étude des différences phénotypiques vis-à-vis du statut de la maladie (appelé discordance) entre les jumeaux MZ refléterait uniquement l'effet de l'environnement sur le phénotype (car les jumeaux MZ partagent le même patrimoine génétique) alors que la discordance chez les jumeaux DZ serait due aux effets environnementaux et génétiques (les jumeaux DZ ne partagent que la moitié du patrimoine génétique). L'héritabilité génétique est estimée par la différence entre le taux de concordance chez les deux types de jumeaux (Falconer & Mackay, 1960) :  $H^2 = 2(r_{MZ} - r_{DZ})$ .

Dans le cas d'un phénotype quantitatif, différentes méthodes ont été développées pour effectuer des études de jumeaux. Ces différentes méthodes sont décrites dans la revue *Classical Twin Studies and Beyond* par Boomsma *et al* en 2002 (Boomsma, Busjahn, & Peltonen, 2002) et ne seront pas détaillées ici.

Les limites des études de jumeaux sont la difficulté de recruter des cohortes de grande taille, le biais de sélection des paires atteintes et le fait que les jumeaux MZ partagent souvent plus leur environnement que les jumeaux DZ (Segal, 2000).

#### 1.1.4 Analyses de ségrégation

Lorsqu'une agrégation familiale est mise en évidence, l'analyse de ségrégation permet de rechercher la présence d'un gène majeur, c'est-à-dire qui a un effet détectable, parmi l'ensemble des facteurs génétiques et environnementaux déterminant le phénotype étudié. Les modèles régressifs proposés par Bonney permettent de rechercher l'effet d'un gène majeur en présence de différentes sources de variation familiale, génétique et/ou environnementale et d'autres facteurs de risque (Bonney, 1984, 1986). Les modèles régressifs sont basés sur les principes de la régression logistique pour un trait binaire (Bonney, 1986) et de la régression linéaire pour un trait quantitatif (Bonney, 1984). Ces modèles permettent d'estimer simultanément les effets des différents facteurs sous différents modèles. Ils sont ensuite comparés par un test de rapport de vraisemblance, permettant de tester la présence de différentes composantes (dépendances familiales, gènes majeurs, facteurs de risque ...). Ces approches sont très puissantes pour l'étude de maladies monogéniques ou dans le cas de maladies multifactorielles avec l'implication d'un ou deux gène(s) à effet fort. Cependant, elles sont peu puissantes dans le cas de l'étude de maladies complexes impliquant l'effet de nombreux gènes à effets faibles pouvant être en interaction entre eux et avec des facteurs d'environnement.

Il existe deux approches complémentaires pour localiser les facteurs génétiques impliqués dans les maladies complexes : les analyses de liaison génétique, qui s'intéressent à la transmission des allèles des *loci* au niveau familial et les analyses d'association génétique, qui s'intéressent à l'association entre les allèles du gène responsable de la maladie et les allèles du marqueur génétique au niveau populationnel.

### 1.1.5 Analyses de liaison génétique

Les analyses de liaison génétique sont fondées sur la ségrégation conjointe du *locus* de la maladie et des marqueurs génétiques. Elles sont basées sur le taux de recombinaison.

#### 1.1.5.1 Le taux de recombinaison

La recombinaison est un phénomène résultant du mélange de matériel génétique qui se produit lors d'évènement inter-chromosomique (*crossing-over*) lors de la méiose<sup>4</sup>. Si l'on considère deux *loci* A et B comportant chacun deux allèles différents : A/a et B/b. Si nous considérons un parent porteur des allèles A et B sur un chromosome et a et b sur l'autre, il pourra transmettre à son enfant quatre types de gamètes : AB et ab, dits gamètes parentaux et Ab et aB, dits gamètes recombinés (Figure 5).

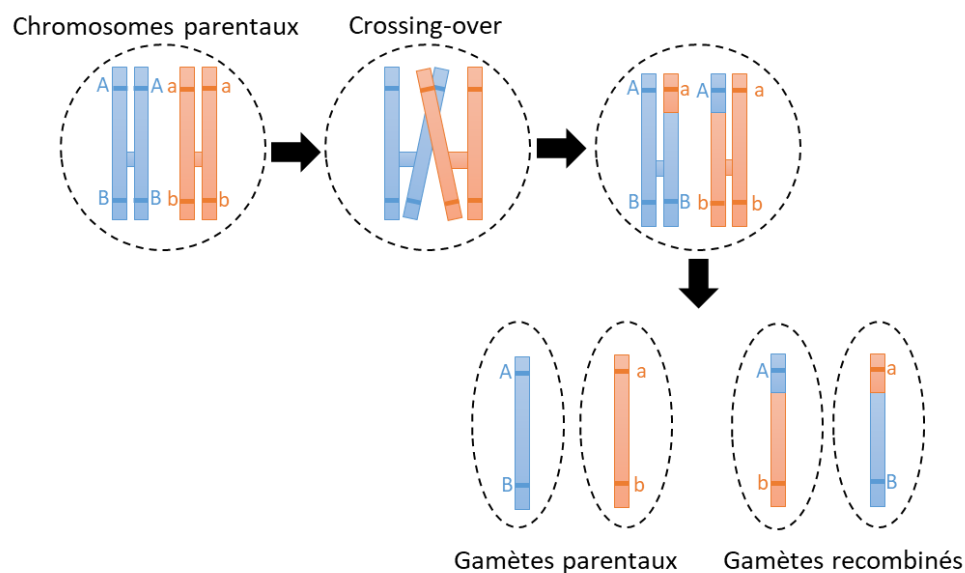


Figure 5 : Crossing-over et gamètes

Le taux de recombinaison  $\theta$  entre deux *locus* A et B correspond à la proportion de gamètes recombinés (ici aB et Ab) parmi l'ensemble des gamètes transmis :

<sup>4</sup> La méiose correspond à un processus biologique qui permet la création des gamètes, qui sont des cellules haploïdes (un seul exemplaire de chaque chromosome), à partir des cellules diploïdes (une paire chromosomes)

$$\theta = \frac{\text{nombre de gamètes recombinés transmis}}{\text{nombre de gamètes total transmis}}$$

Si les deux *loci* sont situés sur deux chromosomes différents ou très éloignés sur le même chromosome, la probabilité d'avoir une recombinaison entre les *loci* est importante. La proportion de gamètes parentaux est alors égale à la proportion de gamètes recombinés et  $\theta$  vaut  $\frac{1}{2}$ . Ces deux *loci* sont alors dits « indépendants ».

Si les deux *loci* sont proches, la proportion de gamètes parentaux est supérieure à la proportion de gamètes recombinés : les deux *loci* sont dit « génétiquement liés » et  $\theta$  est inférieur à  $\frac{1}{2}$ . Plus deux *loci* sont proches et plus la probabilité d'observer un évènement de recombinaison diminue, donc plus la valeur de  $\theta$  est proche de 0. On représente la probabilité de recombinaison sur une distance donnée par l'unité de distance génétique : le centiMorgan (cM), 1 cM correspondant à environ 1% de recombinaison (soit une recombinaison pour 100 méioses).

#### 1.1.5.2 Les différents types d'analyses de liaison génétique

Les tests de liaison génétique permettent de tester si un gène causant une maladie et un marqueur génétique (dont la localisation est connue) co-ségrègent de manière indépendante ou non au cours de la méiose. Il existe deux catégories d'analyse de liaison : les analyses modèle-dépendantes, qui modélisent le mode de transmission du phénotype, et les analyses modèle-indépendantes, qui ne nécessitent pas cette modélisation.

Soit un *locus* biallélique A/a, avec a l'allèle à risque. Quatre modèles de transmission génétique ont été définis, notamment par Gregor Mendel en 1866 (Figure 6) :

- Le modèle additif : lorsque que l'expression du phénotype est proportionnelle au nombre d'allèles à risque (effet dose de l'allèle à risque) ;
- Le modèle dominant : lorsqu'une seule copie de l'allèle à risque suffit pour exprimer le phénotype étudié ;
- Le modèle récessif lorsque les deux copies de l'allèle à risque sont nécessaires pour exprimer le phénotype étudié ;
- Le modèle d'avantage aux hétérozygotes ou d'*over-dominance* où seuls les sujets hétérozygotes (Aa) exprimeront le phénotype.



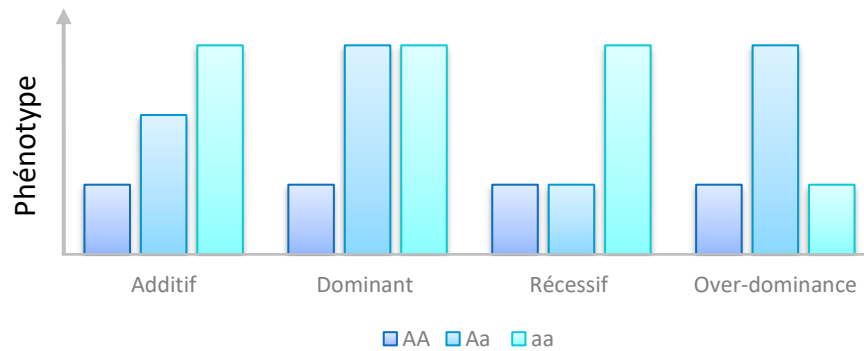


Figure 6 : Représentation d'un phénotype quantitatif pour les quatre principaux modes de transmission

Le test de liaison le plus utilisé est le test du *LOD-Score* (*logarithm of odds*, logarithme des probabilités en anglais), proposé par Morton en 1955 et correspond à une méthode modèle-dépendante. Pour une famille, le rapport de vraisemblance est calculé entre une hypothèse de liaison entre deux *loci* (marqueur étudié et maladie) pour un  $\theta_1$  donné, inférieur à  $\frac{1}{2}$ , et l'hypothèse nulle ( $H_0$ ) d'absence de liaison avec un  $\theta$  de  $\frac{1}{2}$  (Morton, 1955). La statistique du rapport de vraisemblance en  $\theta_1$  de la famille  $i$  est définie par :

$$z_i(\theta_1) = \log_{10} \left( \frac{Li(\theta = \theta_1)}{Li\left(\theta = \frac{1}{2}\right)} \right)$$

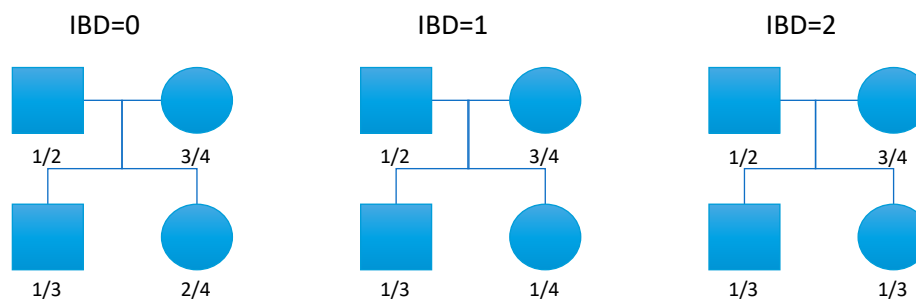
Le LOD-score correspond à la somme pour un nombre de familles donné des rapports de vraisemblance calculés pour chacune de ces familles. La vraisemblance est calculée en fonction du mode de transmission et des paramètres du gène de la maladie (fréquence du gène, pénétrance pour les trois génotypes, ...). Morton a proposé les seuils de décision suivants dans son article princeps :

- Si  $Z(\theta_1) > 3$  :  $H_0$  est rejetée ; les *loci* sont considérés comme génétiquement liés en  $\theta = \theta_1$  ;
- Si  $Z(\theta_1) \leq -2$  :  $H_0$  n'est pas rejetée ; on exclut la liaison génétique en  $\theta_1$ . Les deux *loci* sont considérés comme étant indépendants ;
- Si  $-2 < Z(\theta_1) \leq 3$  : on ne peut pas conclure.

En 1995, Lander a redéfini ces seuils dans le cas des analyses sur l'ensemble du génome, en considérant un seuil de significativité pour une liaison génétique de 3,6 au lieu de 3 (Lander & Kruglyak, 1995).

Les méthodes d'analyse de liaison modèle-dépendantes, comme la méthode du LOD-Score, sont puissantes pour localiser les gènes lorsque le modèle génétique est connu. L'utilisation de ces méthodes est cependant délicate en absence de connaissance du modèle génétique, une mauvaise spécification de ce modèle pouvant conduire à exclure faussement une région du génome (Clerget-Darpoux, Bonaiti-Pellie, & Hochez, 1986).

Dans ce cas-là, il est possible d'utiliser les méthodes modèles-indépendantes. Le principe de ces méthodes est d'étudier la corrélation entre la ressemblance entre les apparentés et le nombre d'allèles reçus identiques par descendance mendélienne ou IBD (pour *Identical By Descent* en anglais). Deux allèles sont IBD lorsqu'ils sont hérités du même ancêtre commun. Le plus souvent, ces méthodes sont basées sur les paires de germains (frères/sœurs, *sib-pair* en anglais) et permettent de tester si les germains, qui se ressemblent pour le phénotype étudié, ont hérité de leurs parents des copies identiques du marqueur plus souvent que ne le voudrait le hasard. Pour illustration, un exemple de paires partageant 0, 1 ou 2 allèles IBD est présenté dans la Figure 7.



**Figure 7 :** Proportion d'allèles partagés pour les différents statuts IBD  
Chaque allèle est représenté par un numéro allant de 1 à 4.

Historiquement, la première étude de paires de germains a été proposée par Penrose (Penrose, 1953), puis différentes variantes de cette méthode ont été ensuite proposées (Shah & Green, 1994), la méthode d'Haseman-Elston étant la plus utilisée (Haseman & Elston, 1972). L'observation des génotypes au *locus* marqueur chez les parents et les enfants des familles permet d'estimer  $\pi_j$ , la proportion d'allèles identiques chez les germains de la paire  $j$ . Sous  $H_0$ ,

les deux *loci* sont considérés comme indépendants, et la probabilité  $\pi_j$  que la paire de germains partage 0, 1 ou 2 allèles IBD sont respectivement de  $\frac{1}{4}$ ,  $\frac{1}{2}$  et  $\frac{1}{4}$ .

La méthode par IBD nécessite de pouvoir identifier l'état IBD des allèles, ce qui est compliqué lorsque des marqueurs génétiques sont peu polymorphes ou lorsque les deux parents ne sont pas génotypés (Lange, 1986). Il existe d'autres méthodes basées sur l'identité par état (ou *Identical By State* ou IBS en anglais), ne nécessitant pas de connaître l'origine parentale de l'allèle partagé, mais elles sont fortement dépendantes de l'estimation des fréquences alléliques (Bishop & Williamson, 1990; Lange, 1986).

## 1.2 Etudes d'association génétique

Historiquement, les analyses d'association permettent d'identifier des variants génétiques impliqués dans la variabilité d'un phénotype à partir de données populationnelles. Plus récemment, des méthodes ont été développées pour les données familiales. Elles sont basées sur le déséquilibre de liaison.

### 1.2.1 L'équilibre d'Hardy Weinberg

Soit un polymorphisme bi-allélique avec  $a$  l'allèle à risque minoritaire (de fréquence  $p_a$ ) et  $A$  le second allèle (de fréquence  $p_A$ ). L'équilibre d'Hardy-Weinberg (HW) est un principe fondamental en génétique, qui définit que les fréquences génotypiques à un locus donné sont constantes d'une génération à l'autre, sous les conditions suivantes :

- La population est de taille infinie ;
- Les unions sont aléatoires, le choix d'un partenaire ne dépend pas de son génotype (on parle de panmixie) ;
- Il n'y a ni sélection, ni mutation, ni migration dans la population.

Sous l'hypothèse d'équilibre d'HW, les fréquences ( $f$ ) génotypiques pour les génotypes  $AA$ ,  $Aa$  et  $aa$  sont respectivement  $f_{AA} = p_A^2$  ;  $f_{Aa} = 2p_A p_a$  et  $f_{aa} = p_a^2$ . Un écart à l'équilibre d'HW peut s'expliquer par le non-respect des conditions d'application de l'équilibre, notamment la présence de mutation, de migration ou de sélection dans les populations. La fréquence de l'allèle minoritaire  $p_a$  sera par la suite appelée MAF pour l'anglais *Minor Allele Frequency*.

L'équilibre d'HW est un point important des analyses d'association. En effet, un écart trop important à cet équilibre peut être dû à des erreurs de génotypage, la vérification de l'écart à l'équilibre d'HW est donc un des points du contrôle qualité effectué avant le GWAS (Marees *et al.*, 2018; Reed *et al.*, 2015). Cependant, cet écart à l'équilibre d'HW peut aussi être dû à une association entre le marqueur et le trait étudié. Dans le cas d'étude cas-témoins, ce contrôle qualité est ainsi effectué uniquement chez les témoins, ou avec un seuil moins strict chez les cas (Marees *et al.*, 2018; Reed *et al.*, 2015).

### 1.2.2 Le déséquilibre de liaison

Les analyses d'association sont basées sur le déséquilibre de liaison. On dit que deux *loci* sont en déséquilibre de liaison (ou en LD pour *Linkage Disequilibrium* en anglais) lorsqu'une combinaison spécifique des allèles de ces *loci* est plus fréquente que les combinaisons attendues aléatoirement.

Soit deux marqueurs bi-alléliques, A/a et B/b, de fréquences  $p_A/p_a$  et  $p_B/p_b$  respectivement. Il y a donc quatre haplotypes<sup>5</sup> différents, identiques aux gamètes présentés précédemment (Chapitre 1.1.5.1) : un haplotype comprenant les allèles A et B, un autre comprenant les allèles A et b, un troisième haplotype comprenant les allèles a et B, et enfin un dernier haplotype comprenant les allèles a et b. Le déséquilibre de liaison est estimé par le paramètre D défini par :  $D = p_{AB} - p_A p_B$  avec  $p_{AB}$  la fréquence des haplotypes porteurs de A et de B. On parle alors de l'haplotype AB. En absence de déséquilibre de liaison, les allèles des deux *loci* se répartissent de façon aléatoire dans les gamètes et D est égal à 0, et les quatre haplotypes sont présents dans des répartitions identiques dans la population. En présence de déséquilibre de liaison, D est différent de 0 et l'un des haplotypes est transmis de façon préférentielle dans la population. La valeur de D ne permet pas de quantifier le déséquilibre de liaison, ni de comparer les valeurs de D entre différents *loci*. Des standardisations de D ont été proposées, dont les plus utilisées sont  $D'$  (Lewontin, 1964, 1988) et  $r^2$  (Hill & Robertson, 1968).

---

<sup>5</sup> Un haplotype correspond à une combinaison d'allèle transmis conjointement.

La première standardisation proposée,  $D'$ , consiste à standardiser la valeur de  $D$  par la valeur maximale qu'il peut prendre :

$$D' = D/D_{max} \text{ ou } D_{max} = \begin{cases} \min\{p_A p_b; p_a p_B\} & \text{si } D > 0 \\ \min\{p_A p_B; p_a p_b\} & \text{si } D < 0 \end{cases}$$

$D'$  est compris entre 1 et -1. Les valeurs 1 et -1 signifient que l'un des quatre haplotypes n'est pas observé. En absence de déséquilibre de liaison, de la même manière que  $D$ ,  $D'$  est égal à 0.

La seconde standardisation,  $r^2$ , consiste à calculer le carré du coefficient de corrélation entre les deux *loci* étudiés (Hill & Robertson, 1968) :

$$r^2 = \frac{D^2}{p_A p_a p_B p_b}$$

La valeur de  $r^2$  varie entre 0 et 1, 1 correspondant à un déséquilibre complet, c'est-à-dire que seuls deux haplotypes sont observés. Un seuil de 0.8 est utilisé pour définir un fort déséquilibre de liaison (Carlson *et al.*, 2004). Cependant un seuil différent peut être choisi en fonction des fréquences alléliques des marqueurs étudiés (Wray, 2005).

Des blocs de déséquilibre de liaison, ou bloc haplotypique, vont ainsi être définis par le  $D'$  ou le  $r^2$ , correspondant à un groupe d'allèle de différents *loci* situés sur un même chromosome et transmis conjointement. L'identification des blocs haplotypiques a permis de rendre possible l'analyse du génome entier à partir de l'analyse de quelques centaines de milliers de SNPs, appelés *tagSNPs*, couvrant la majorité du génome (Li, Li, & Guan, 2008).

Les analyses d'association sont sensibles à la stratification de population (Marchini, Cardon, Phillips, & Donnelly, 2004). En effet, le projet Hapmap a démontré que les blocs de déséquilibre de liaison pouvaient être de tailles différentes selon les populations humaines (Frazer, Murray, Schork, & Topol, 2009). Différentes stratégies ont été développées pour prendre en compte la présence de stratification de population dans les analyses d'association, la plus utilisée étant la méthode d'analyse en composante principale (PCA pour *Principal Components Analysis*) (Price *et al.*, 2006). Cette méthode consiste, dans le cadre de données génétiques, à définir des composantes, permettant de définir des populations à partir des données génotypiques.

### 1.2.3 Le principe des études d'association génétiques

#### 1.2.3.1 Méthodes d'association génétique

Les premières analyses d'association pan-génomiques haut débit ont été utilisées pour l'analyse de traits qualitatifs (étude cas-témoins) (Clarke *et al.*, 2011), dont la première a été publiée en 2002 (Ozaki *et al.*, 2002). Depuis, des études d'association sur des traits quantitatifs ont aussi été développées. Dans les deux cas, les méthodes classiques d'analyses d'association consistent à étudier indépendamment chacun des marqueurs choisis, en effectuant autant de tests statistiques que de marqueurs étudiés (De, Bush, & Moore, 2014). D'autres stratégies consistant à étudier conjointement plusieurs marqueurs ou directement des haplotypes ont été développées, mais ne seront pas présentées dans ce travail.

Les études d'association peuvent s'effectuer sur un groupe de régions sélectionnées au préalable. Il s'agit alors d'une stratégie gènes-candidats (Kwon & Goate, 2000), ou sur l'ensemble du génome, et il s'agit alors d'étude pan-génomique ou *Genome-wide Association Study* (GWAS) (Bush & Moore, 2012).

La stratégie gène-candidat consistent à sélectionner un(e) ou plusieurs gènes ou régions, à partir de données ou connaissances *a priori* de la maladie ou du trait étudié et de tester l'association génétique. L'étape cruciale de ce type de stratégie consiste à identifier et à sélectionner les marqueurs qui seront testés (Kwon & Goate, 2000). Un point important de ce type d'analyse est le nombre important de faux positifs, qui peuvent être notamment dû à des études effectuées sur des échantillons particuliers, avec par exemple une stratification de population importante entre les cas et les témoins (Sullivan, 2007). Pour cela il est donc important de toujours bien définir des caractéristiques précises pour le trait étudié, et notamment dans les cohortes de répliation.

La stratégie génome-entier consiste à étudier un ensemble de marqueurs génétiques, généralement des *tagSNPs*, permettant d'analyser l'ensemble du génome. Pour être considérée comme significativement associée, une région d'association (définie par la région de LD du SNP associé) doit avoir été répliquée dans une cohorte indépendante (McCarthy *et al.*, 2008). Il est possible d'effectuer une stratégie à deux étapes (Skol, Scott, Abecasis, & Boehnke, 2006), la première étape consiste en un GWAS, qui permet de sélectionner les régions les plus

significativement associées au trait dans un premier échantillon. Ces marqueurs seront ensuite testés dans un deuxième échantillon, considéré comme échantillon de réplication. Cette stratégie permet d'augmenter la puissance du test d'association en diminuant le nombre de tests.

Actuellement, il existe plusieurs bases de données référençant les résultats de la majorité des GWAS effectués sur des phénotypes chez l'Homme (aussi bien quantitatifs, comme le taux d'HbA1c, que binaires, comme les études cas/témoins pour le développement du DT1), comme GWAS Catalog (MacArthur *et al.*, 2017) ([www.ebi.ac.uk/gwas](http://www.ebi.ac.uk/gwas)), référençant les marqueurs significativement associés au phénotype étudié et leurs caractéristiques ainsi que les résultats statistiques du GWAS correspondant.

Une des problématiques d'étudier chaque marqueur de façon indépendante est de prendre en compte le nombre de tests, qui peut être supérieur à un million pour certain GWAS, correspondant au nombre de SNPs analysés présents sur la puce. Il est important aussi de souligner que chacun des marqueurs testés n'est pas totalement indépendants des autres marqueurs présents sur la puce, notamment à cause de la présence du déséquilibre de liaison (LD) entre les marqueurs. Différentes méthodes ont été proposées comme la correction de Bonferroni ou le FDR (pour *False discovery rate*). La correction de Bonferroni consiste à considérer comme seuil de significativité le risque de première espèce divisé par le nombre de tests effectués (ici le nombre de SNPs analysés  $k$ ) :  $\alpha_{Cor} = \frac{\alpha}{k}$ . S'il existe un déséquilibre de liaison entre plusieurs SNPs testés, les tests d'association effectués ne seront pas totalement indépendants et la correction de Bonferroni sera trop conservatrice. Des méthodes ont été développées pour estimer le nombre réel de tests indépendants, dont la méthode de Gao (Gao, Starmer, & Martin, 2008). À l'aide d'une analyse en composantes principales, les SNPs vont être regroupés en fonction de leur LD, et ainsi permettre d'estimer le nombre de tests indépendants. La correction de Bonferroni est ensuite appliquée sur ce nombre de tests.

Des méthodes de permutation peuvent aussi être utilisées pour définir le seuil de significativité, mais elles nécessitent un temps de calcul important et sont très conservatrices en présence de fort LD (Gao, 2011).

Le FDR correspond à la proportion d'erreurs commises en rejetant à tort l'hypothèse nulle parmi le nombre d'hypothèses nulles rejetées (Benjamini & Hochberg, 1995) mais est peu utilisé dans le cadre des études génétiques de type GWAS (Sham & Purcell, 2014).

Dans le cadre de l'étude d'un trait quantitatif, le seuil de significativité consensuel est de  $5 \cdot 10^{-8}$ , correspondant à la correction de Bonferroni appliqué sur un nombre théorique d'un million de tests (Risch & Merikangas, 1996).

Un des autres points cruciaux des études d'association est la caractérisation du phénotype étudié. En effet, dans le cas de traits complexes, il peut exister une variabilité qui peut entraîner une diminution de la puissance statistique (Bush & Moore, 2012), aussi bien dans la caractérisation des malades et des témoins pour les traits binaires que pour les méthodes de mesure dans le cas des traits quantitatifs.

Avant d'effectuer les tests d'association, il est important de vérifier plusieurs points en effectuant un contrôle qualité pour que les résultats significatifs ne soient pas imputables à des erreurs au moment du génotypage (contamination, mauvaises amorces, etc.) ou de mauvaises qualité du jeu de données (ADN trop vieux, erreur de tube, etc.). Ce contrôle qualité consiste à filtrer les individus et les marqueurs selon plusieurs critères avec des seuils qui dépendront de chaque analyse : (1) le taux de génotypage de chaque marqueur et de chaque individu ; (2) la correspondance entre le sexe phénotypique et génotypique ; (3) la fréquence de l'allèle mineur (MAF) ; (4) l'écart à l'équilibre d'HW ; (5) le taux d'hétérozygotie des marqueurs ; (6) le lien de parenté entre les individus et (7) la stratification de population. Un tutoriel avec les différentes étapes et seuils conseillés a été publié récemment (Marees *et al.*, 2018).

Enfin, le ou les marqueurs associés au trait étudié sont rarement le variant causal, mais sont plus généralement en LD avec celui-ci, et il est alors question d'association indirecte (Hirschhorn & Daly, 2005). Cela permet d'identifier des régions associées à des phénotypes sans pour autant avoir besoin de génotyper l'ensemble des marqueurs du génome, cependant cela entraîne aussi l'obligation d'effectuer des analyses supplémentaires afin d'identifier le marqueur causal (Bush & Moore, 2012).

Une fois l'association détectée, il est possible de mesurer la force de l'association entre le trait et le génotype. Pour cela, dans le cas d'un trait binaire, elle est mesurée à l'aide du calcul d'un



*odds-ratio* (OR) (Szumilas, 2010). Pour un trait quantitatif, la force de l'association peut être estimée par le pourcentage de variance expliqué (PEV) ou par la taille de l'effet, mesurés par le coefficient de régression ( $\beta$ ). Dans le cas d'une méthode basée sur une régression univariée, le PEV est estimé à partir du coefficient de détermination  $R^2$ , qui correspond au rapport de la variance expliquée par le facteur sur la variance totale.

### 1.2.3.2 Les méthodes d'analyses d'association pour l'étude d'un trait quantitatif

Différentes méthodes ont été développées, qui peuvent être modèles-dépendantes, c'est-à-dire nécessiter de modéliser préalablement le mode de transmission du marqueur, ou modèles-indépendantes, ne nécessitant pas de modélisation.

#### Les méthodes modèles-dépendantes

Les méthodes d'analyse d'association les plus utilisées sont les méthodes basées sur la régression linéaire (LR pour *Linear Regression* en anglais) modélisant le phénotype à l'aide d'une variable sous un mode de transmission particulier. Le modèle de régression s'écrit :

$$P = \alpha + \beta_{mod} \cdot X_{mod} + \beta_{COV} \cdot COV + E$$

avec P le phénotype quantitatif,  $X_{MOD}$  le vecteur de génotype en fonction du modèle génétique choisi. COV représente la matrice des covariables et E la matrice des effets aléatoires.

Le codage des génotypes en fonction du modèle est donné dans le Tableau 1 (MOD = ADD pour additif, DOM pour dominant, REC pour récessif et HET pour un avantage aux hétérozygotes).

Modèle	Additif LR_ADD	Dominant LR_DOM	Récessif LR_REC	Over dominance LR_HET
Génotypes	$X_{add}$	$X_{dom}$	$X_{rec}$	$X_{het}$
AA	0	0	0	0
Aa	1	1	0	1
aa	2	1	1	0

**Tableau 1 :** Codage du génotype en fonction du modèle génétique considéré

Sous l'hypothèse d'absence d'association,  $\beta_{mod}$  est nul. La statistique  $Z_{mod}$  est calculée à l'aide de la formule suivante et qui suit sous  $H_0$  une distribution de  $\chi^2$  à un degré de liberté :

$$Z_{mod} = \frac{\beta_{mod}^2}{var(\beta_{mod})}$$

Lors d'études à haut débit et afin de ne pas augmenter le nombre de tests, en testant l'association pour chacun des modèles, le LR\_ADD est classiquement utilisé, étant considéré comme un bon compromis entre les différentes modélisations des trois modes de transmission classiques (Balding, 2006). Comme pour le LOD-Score, les méthodes d'association modèle-dépendantes ont pour avantage d'être très puissantes lorsque l'on connaît le mode de transmission du trait étudié.

### Les méthodes modèles-indépendantes

#### *Modélisation du modèle génétique à l'aide de deux variables indicatrices*

Une première approche modèles-indépendantes consiste à modéliser les génotypes à l'aide de plusieurs variables indicatrices. Le codage de ces variables va ainsi permettre de tester l'association sans faire d'hypothèse sur le mode de transmission. PLINK propose deux codages différents (Tableau 2) (Chang *et al.*, 2015; Purcell *et al.*, 2007). Le GRM, dont l'un de mes travaux de thèse consistait à étudier ses propriétés statistiques, correspond au premier codage et est appelé LR\_GENO dans PLINK.

Le modèle s'écrit :

$$P = \alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_{COV} \cdot COV + E$$

avec P le phénotype quantitatif,  $X_1$  et  $X_2$  les deux vecteurs de génotype dont les codages sont présents dans le Tableau 2. COV représente la matrice des covariables et E la matrice des effets aléatoires.

	GRM ou LR_GENO		LR_HETHOM	
Génotypes	$X_1$	$X_2$	$X_1$	$X_2$
AA	0	0	0	0
Aa	1	1	1	0
aa	2	0	0	1

**Tableau 2 :** Codage des variables pour les tests génotypiques disponibles sur PLINK (Chang *et al.*, 2015; Purcell *et al.*, 2007)

Il est possible de tester l'association à l'aide d'un test de Fisher, mais la majorité des études étant effectuée à l'aide du logiciel PLINK, le test effectué correspond à un test de rapport de vraisemblance. Dans le cas d'une étude sans covariables, dans PLINK, sous l'hypothèse  $H_0$  d'absence d'association,  $\beta_1 = \beta_2 = 0$ , le test est donc un test de rapport de vraisemblance entre

les modèles M1 sans variable indicatrice et le modèle M2 avec les deux variables, qui sous H0 suit une loi de  $\chi^2$  à 2 degrés de liberté (Chang *et al.*, 2015; Purcell *et al.*, 2007) :

- M<sub>1</sub> : P =  $\alpha$ , de vraisemblance  $L(M_1)$
- M<sub>2</sub> : P =  $\alpha + \beta_2.X_1 + \beta_1.X_2$ , de vraisemblance  $L(M_2)$ .

Le test s'écrit alors :

$$LR = -2 \times \ln \frac{L(M_1)}{L(M_2)}$$

Une des critiques de ce test est qu'il possède deux degrés de liberté, décrit comme entraînant une perte de puissance ainsi qu'une plus grande complexité d'interprétation (Balding, 2006).

L'étude de Lette *et al* a comparé la différence de puissance des différents tests d'association génétique basés sur la régression linéaire, c'est-à-dire trois des LR modèles-dépendants, des combinaisons de ces trois tests ainsi que le LR\_GENO (Lette, Lange, & Hirschhorn, 2007). Pour cela, les auteurs ont simulé un trait continu centré réduit pour un jeu de données de 1000 individus. Les marqueurs associés à ce trait ont été simulés avec une MAF variant entre 5 et 50%, les trois modes de transmission (additif, récessif, et dominant) et un PEV de 1%. Pour chaque MAF et mode de transmission, 10 000 simulations ont été analysées avec les trois LR (LR\_ADD, LR\_DOM, et LR\_REC), une combinaison d'au moins deux LR (LR\_ADD ou LR\_DOM associés à LR\_REC, ou les trois LR) et le LR\_GENO. Dans le cas des combinaisons, deux types de corrections pour les tests multiples ont été appliqués, basés sur des permutations ou en utilisant la correction de Bonferroni.

Lette *et al* ont ainsi pu mettre en évidence une perte de puissance des LR modèles-dépendants lorsque le mode de transmission est mal spécifié (Lette *et al.*, 2007). Il existe une perte de puissance en utilisant le LR\_ADD lorsqu'il existe un écart à l'additivité (pour un modèle récessif ou dominant par exemple). Les tests LR\_ADD et LR\_DOM sont aussi puissants sous un modèle additif ou dominant, excepté quand la MAF est supérieure à 0,4. Dans le cas d'un modèle simulé récessif, le test LR\_ADD est moins puissant que le LR\_REC (jusqu'à 60% pour une MAF de 0,05).

La puissance de la combinaison des trois tests est un bon compromis sous les trois modèles génétiques en estimant le risque de première espèce par permutation mais le temps de calcul

est assez long et s'adapte mal aux études pangénomiques. L'application de la correction de Bonferroni, plus rapide, rend le test très conservateur (Lettre *et al.*, 2007).

La puissance du LR\_GENO a aussi été comparée aux régressions linéaires à un degré de liberté, et montre que le LR\_GENO est pratiquement aussi puissant que la régression linéaire correspondant au modèle simulé, et globalement plus puissant que les régressions linéaires sous les autres modes de transmission. Lorsque la puissance du LR\_GENO est comparée à celle des combinaisons de régression linéaire à un degré de liberté, le LR\_GENO est aussi puissant que la meilleure combinaison pour chacun des cas représentés. Cette étude recommandait d'utiliser soit le LR\_GENO soit la combinaison de l'ensemble des tests, avec une correction adéquate pour les tests multiples (Lettre *et al.*, 2007).

Cependant, l'ensemble de cette étude a été effectué en ne testant qu'un seul PEV de 1%, il est donc nécessaire de valider ces résultats sur un plus grand nombre de PEV.

#### *Les méthodes optimisées ou tests MAX*

Le mode de transmission étant généralement inconnu, il est possible de tester l'association en utilisant le maximum de la version standardisée des tests optimaux (MAX pour *Maximum of the standardized optimal tests*) (Freidlin, Zheng, Li, & Gastwirth, 2002). Plusieurs tests MAX ont récemment été développés pour l'analyse des traits quantitatifs, dont un seulement sera présenté ici (Loley, König, Hothorn, & Ziegler, 2013). Dans cette version, les génotypes AA, Aa et aa sont codés à l'aide de 3 variables indicatrices,  $X_{AA}$ ,  $X_{Aa}$  et  $X_{aa}$  (Tableau 3). Le modèle s'écrit alors :

$$P = \beta_{AA} \cdot X_{AA} + \beta_{Aa} \cdot X_{Aa} + \beta_{aa} \cdot X_{aa} + \beta_{COV} \cdot COV$$

avec P le phénotype quantitatif,  $X_j$  les vecteurs de génotype dont les codages sont présents dans le Tableau 3 (pour j valant AA, Aa et aa). COV représente la matrice des covariables.

	$X_{AA}$	$X_{Aa}$	$X_{aa}$
AA	1	0	0
Aa	0	1	0
aa	0	0	1

**Tableau 3 :** Codage des génotypes

Les paramètres sont estimés par maximum de vraisemblance. Les différents modes de transmission sont modélisés en introduisant des contrastes sur les coefficients de régression  $\beta_j$  (Tableau 4).

	Modèle Additif	Modèle dominant	Modèle récessif
$\beta_{AA}$	-1	-1	$-\frac{1}{2}$
$\beta_{Aa}$	0	$\frac{1}{2}$	$-\frac{1}{2}$
$\beta_{aa}$	1	$\frac{1}{2}$	1
Modèle	$P = -\beta_{AA} + \beta_{aa} + E$	$P = -\beta_{AA} + \frac{1}{2} \beta_{Aa} + \frac{1}{2} \beta_{aa} + E$	$P = -\frac{1}{2} \beta_{AA} - \frac{1}{2} \beta_{Aa} + \beta_{aa} + E$

Tableau 4 : Contrastes des paramètres  $\beta$

La statistique de test, correspondant au maximum de la statistique sous chacun des modèles génétiques s'écrit :

$$T = \max(T_{add}, T_{dom}, T_{rec}) = \max_{j=1}^3 \frac{|\hat{\xi}_j - \xi_j^0|}{\sqrt{[\hat{\Omega}]_{jj}}}$$

Avec  $\hat{\xi} = \xi(\hat{\beta}) = \begin{pmatrix} -\hat{\beta}_{AA} + \frac{1}{2}\hat{\beta}_{Aa} + \frac{1}{2}\hat{\beta}_{aa} \\ -\hat{\beta}_{AA} + \hat{\beta}_{aa} \\ -\frac{1}{2}\hat{\beta}_{AA} - \frac{1}{2}\hat{\beta}_{Aa} + \hat{\beta}_{aa} \end{pmatrix}$  et  $[\hat{\Omega}]_{jj}$  l'estimation de la variance de  $\hat{\xi}_j$

À l'exception d'un mode de transmission simulé récessif, pour lequel la puissance du test MAX est très proche de celles du test LR\_REC, le test MAX est moins puissant que le test LR\_ADD et LR\_DOM pour des MAF inférieurs à 0.3, et est similaire pour des MAF supérieures (Loley *et al.*, 2013).

Le test MAX ne permet pas de tester formellement le modèle génétique. Néanmoins, il est possible de définir le mode de transmission le plus probable à l'aide de la p-valeur qui a été calculée pour chacun des modèles testés. Le modèle le plus probable est ainsi défini comme étant celui dont la valeur de la statistique du test est la maximale ou de façon équivalente celui pour lequel la p-valeur est la plus faible, comme cela avait été définie précédemment pour les tests MAX dans le cadre d'un trait binaire (Hothorn & Hothorn, 2009). Pour un mode de transmission simulé dominant ou récessif, cette démarche permet de retenir le bon mode de transmission, mais dans le cas d'une simulation sous un modèle additif, le mode de transmission généralement retenu est un mode de transmission dominant.

Il est important de noter que ce test MAX demande un temps de calcul important, le temps d'exécution étant au moins doublé pour les petits jeux de données par rapport à un test LR sous un seul modèle.

### *Le test allélique*

Parmi les tests proposés, un test allélique a été proposé en 2013 pour les phénotypes quantitatifs, inspiré du test allélique utilisé dans les études cas-témoins (Lee, Karrison, Cox, & Im, 2013). Ce test permet de comparer les moyennes phénotypiques de deux groupes « Allèle A » et « allèle a », à l'aide d'un t-test ou d'un test de rang non-paramétrique, en fonction du trait étudié et de la taille de l'échantillon. Afin de comparer le test allélique au test LR\_ADD, seule la méthode basée sur un trait centré réduit sera présentée ici.

Pour un phénotype Y quantitatif, centré et réduit, qui est associé à un allèle, il est possible d'attribuer l'effet de l'allèle provenant de la mère, appelé allèle maternel M, et celui provenant du père, l'allèle paternel P :

$$\begin{pmatrix} Y \\ Y \end{pmatrix} = \begin{pmatrix} M \\ P \end{pmatrix} \beta_{all} + \begin{pmatrix} \epsilon_M \\ \epsilon_P \end{pmatrix}$$

Avec M et P respectivement l'allèle maternel et paternel, avec M et P valant 0 ou 1 en fonction de l'allèle à risque, et  $\epsilon_M$  et  $\epsilon_P$  les termes d'erreurs.

Il est alors possible, en dupliquant le phénotype, de tester l'association à l'aide de cette régression linéaire.

Le paramètre  $\beta_{all}$  de la régression allélique sera alors identique au paramètre  $\beta_{add}$  de la régression LR\_ADD. La régression est équivalente à un t-test, puisque chaque variable ne possède que deux catégories. Cependant, la duplication des phénotypes induit une corrélation entre les deux groupes, qu'il est nécessaire de prendre en compte pour tester le modèle. Afin de corriger cette corrélation, il est important de connaître l'importance de l'augmentation de la corrélation sur la variance de  $\beta_{all}$ , en calculant le ratio suivant :

$$\frac{\widehat{var}(\beta_{all} | \text{ignorant la corrélation})}{\widehat{var}(\beta_{all} | \text{en prenant en compte la corrélation})}$$

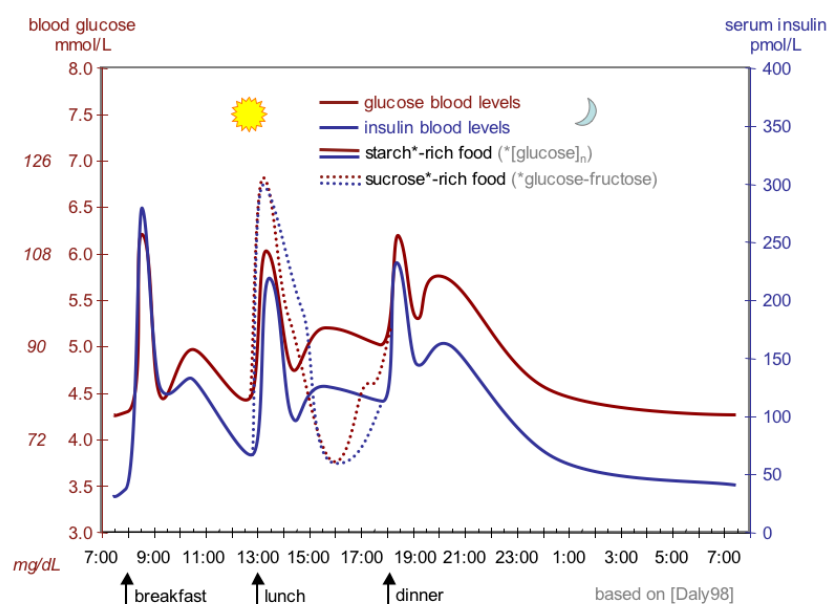
Ce ratio de  $\beta_{all}$  se fait à l'aide du coefficient de corrélation  $R^2$  de la régression, ainsi que la taille de la population.

Une étude de simulation a pu mettre en évidence que le paramètre  $\beta_{add}$  de la régression LR\_ADD est bien identique au paramètre  $\beta_{all}$ , mais que le coefficient de corrélation  $R^2$  sous un modèle allélique est environ deux fois plus important que celui sous un modèle additif. Les études de simulation montrent que ce test allélique a une puissance similaire à celle du LR\_ADD. Cependant, l'étude de simulation effectuée n'a pas pris en compte le mode de transmission, et le test allélique n'a été comparé qu'au test LR\_ADD. Nous ne savons donc pas comment ce test réagit quand le mode de transmission varie. De plus, ce test ne permet pas de prendre en compte des co-variables, ni de tester le modèle génétique sous-jacent.

## 1.3 HbA1c, mesure de l'équilibre glycémique

### 1.3.1 Le taux d'HbA1c : reflet de l'équilibre glycémique sur les trois derniers mois

Le taux de glucose sanguin (ou glycémie) varie énormément au cours du temps, en fonction de l'heure et du type de repas, et notamment de la quantité et du type de sucre ingéré. La glycémie doit être régulée pour un bon fonctionnement de l'organisme, essentiellement grâce à la production et à la sécrétion d'insuline par les cellules bêta des îlots de Langerhans pancréatiques (Figure 8) (Daly *et al.*, 1998).



**Figure 8 :** Variation de la glycémie et du taux d'insuline durant la journée, en fonction du type de repas adapté de Daly *et al.* 1998 (Daly *et al.*, 1998)

Avec *glucose blood level* : glycémie ; *insulin blood level* : taux d'insuline dans le sang ; *starch\*-rich food* : repas riche en amidon ; *sucrose\*-rich food* : repas riche en sucrose ; *breakfast* : petit déjeuner ; *lunch* : déjeuner ; *dinner* : dîner ; *blood glucose* : glycémie et *serum insulin* : insuline sérique.

Chez l'Homme, l'hémoglobine peut prendre plusieurs formes (Allen, A., & Balog, 1958). Dès 1969, il a été montré qu'il existe une surreprésentation d'une de ces formes, appelée hémoglobine glyquée ou HbA1c, chez les personnes souffrant de diabète (Rahbar, Blumenfeld, & Ranney, 1969). L'HbA1c est une glycoprotéine représentant entre 4 et 6% de l'hémoglobine totale chez l'Homme (Bookchin & Gallop, 1968). L'HbA1c est composée de deux chaînes  $\alpha$  et de deux chaînes  $\beta$  sur lesquelles est fixée une molécule de glucose (Figure 9) (Bunn, Haney, Gabbay, & Gallop, 1975; Sherwani, Khan, Ekhzaimy, Masood, & Sakharkar, 2016).



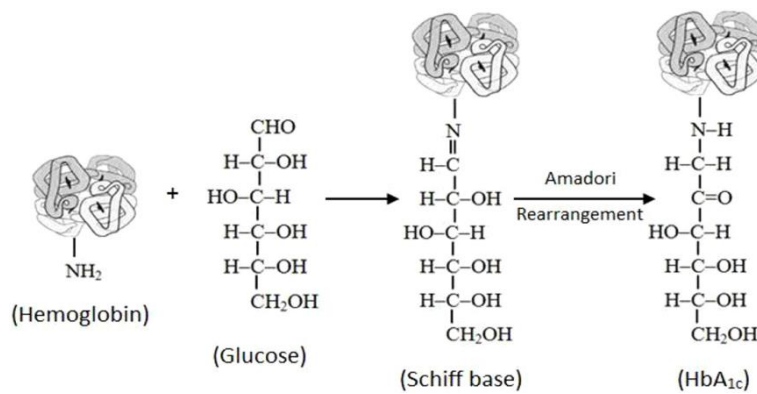


Figure 9 : Structure de l'HbA<sub>1c</sub> – tiré de Sherwani (Sherwani *et al.*, 2016)

L'hémoglobine étant présente uniquement dans les globules rouges (ou érythrocytes) dont la durée de vie est de 120 jours, et le taux de fixation de glucose sur l'hémoglobine étant proportionnelle à la glycémie, le taux d'HbA<sub>1c</sub> reflète la glycémie sur trois mois (Sherwani *et al.*, 2016). Le taux d'HbA<sub>1c</sub> est soit exprimé en pourcentage, comme proposé par DCCT (DCCT, 1994), soit en mmol/mol, comme proposé par l'*International Federation of Clinical Chemistry and Laboratory Medicine* (IFCC) (Comitee, 2007). Le passage d'une unité à l'autre se fait à l'aide de la formule suivante :

$$\text{HbA}_{1c} \text{ (mmol/mol)} = (\text{HbA}_{1c} \text{ (\%)} \times 10.93) - 23.50.$$

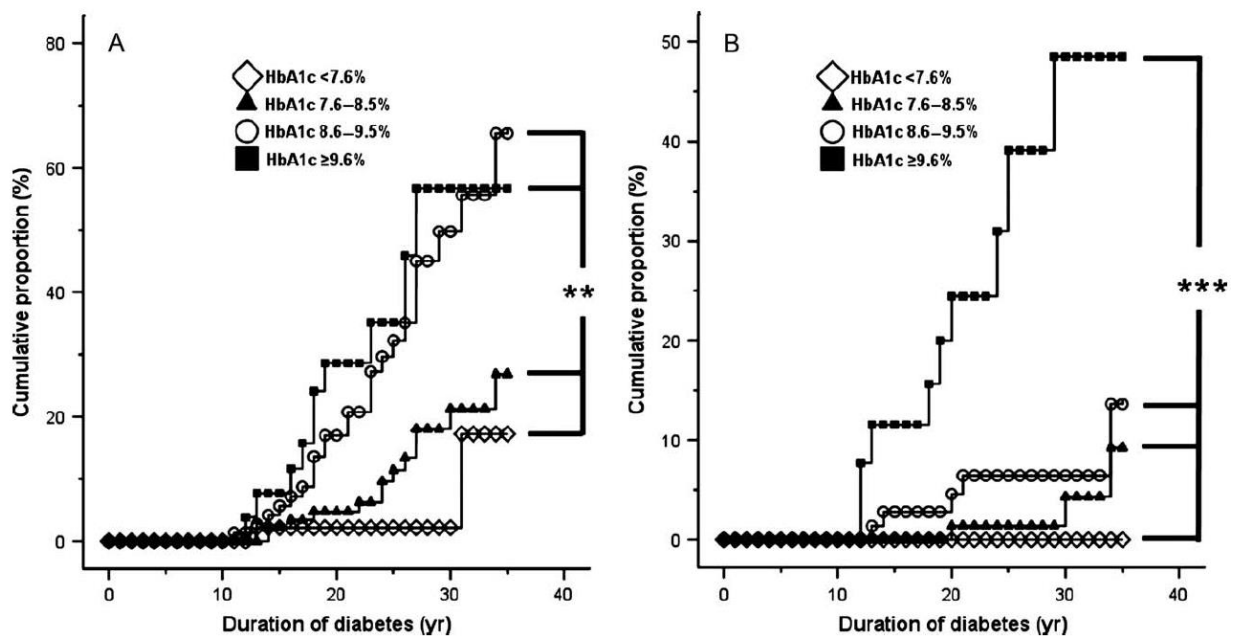
Cliniquement, une variation de 0,5% (5 mmol/mol) du taux d'HbA<sub>1c</sub> est considérée comme significative (Little, Rohlfing, & Sacks, 2011). Pour plus de clarté, les mesures seront toujours présentées en pourcentage dans la suite du manuscrit.

### 1.3.2 Taux d'HbA<sub>1c</sub> et diabète

Le diabète est le plus souvent diagnostiqué à l'aide de la glycémie à jeun. En 2010, le taux d'HbA<sub>1c</sub> a été proposé comme mesure de diagnostic du diabète (ADA, 2011), en considérant un seuil pathologique de 6.5% (Comitee, 2009). Les études de sensibilité et de spécificité ont comparé les différentes méthodes de diagnostic et ont montré que le taux d'HbA<sub>1c</sub> avait une meilleure spécificité que les autres tests glycémiques mais une moins bonne sensibilité (Bennett, Guo, & Dharmage, 2007; Cavagnoli *et al.*, 2011). L'utilisation du seuil de 6.1% permettrait de gagner en sensibilité (Cavagnoli *et al.*, 2011). Cependant, le taux d'HbA<sub>1c</sub> ne peut pas être utilisé comme critère diagnostique chez les patients atteints de certaines maladies

du sang comme les hémoglobinopathies. En pratique, il n'est pas ou peu utilisé pour le diagnostic du DT1 en France, contrairement aux recommandations internationales.

Au cours des années 1990, le taux d'HbA1c a été identifié comme un marqueur du risque de complications aussi bien dans le cas du DT1 que du DT2 (DCCT, 1994; UKPDS, 1998). Plus le taux d'HbA1c est élevé, plus le risque de développer une rétinopathie ou une néphropathie est important (Figure 10) (Andreasson, Ekelund, Landin-Olsson, & Nilsson, 2018; Chalew, McCarter, & Hempe, 2013; Nordwall *et al.*, 2015; Nordwall, Arnqvist, Bojestig, & Ludvigsson, 2009).



**Figure 10 :** Risque de développer une complication en fonction du taux d'HbA1c

Proportion de rétinopathies sévères (A) et de néphropathies déclarées (B) dans une population de DT1 diagnostiquée durant l'enfance en fonction de la durée du DT1, d'après Nordwall *et al* 2009 (Nordwall *et al.*, 2009)

Comparaison entre chaque catégorie du taux d'HbA1c : \*\* p<0.01 ; \*\*\* p<0.001

Afin de limiter la survenue de complications, des recommandations ont été proposées par différentes associations, dont l'ADA, l'*International Society for Pediatric and Adolescent Diabetes* (ISPAD), ou l'Association Canadienne du Diabète (ADA, 2016; Rewers *et al.*, 2014). Dans un premier temps, le seuil du taux d'HbA1c a été défini en fonction de l'âge : inférieur à 8,5% entre 0 et 6 ans ; inférieur à 8% entre 6 et 12 ans et inférieur à 7.5% entre 12 et 18 ans. Cette recommandation avait pour objectif d'éviter la survenue d'hypoglycémies chez les patients très jeunes (ADA, 2011). Depuis 2016, les recommandations de l'ADA ne sont plus fonction de l'âge, le seuil étant fixé à 7.5%, identique à celui recommandé par l'ISPAD (Rewers *et al.*, 2014).

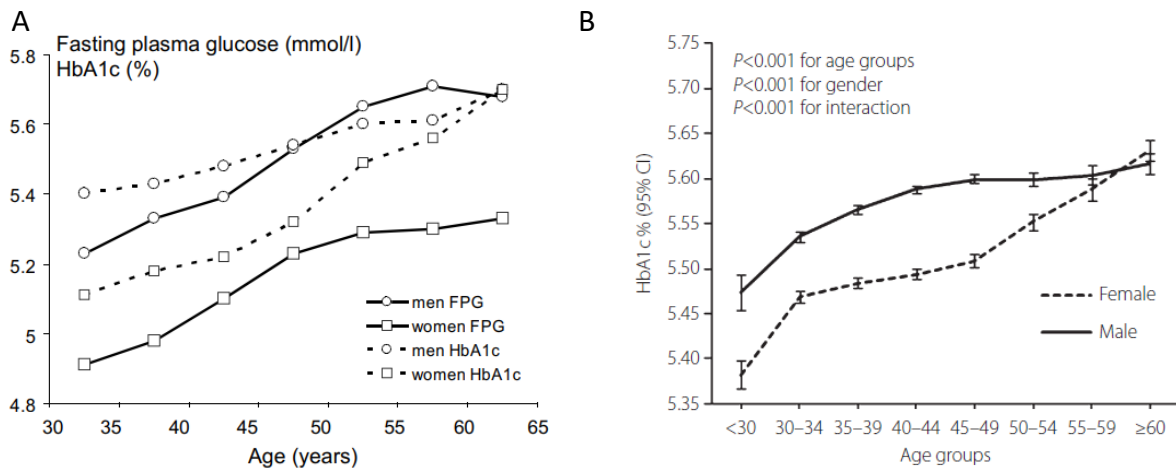
### 1.3.3 Facteurs non-génétiques influençant le taux d'HbA1c chez les personnes non-diabétiques

L'objectif de ma thèse est d'identifier les facteurs génétiques ou non qui influencent le taux d'HbA1c chez l'enfant souffrant de DT1. Cependant, différentes études se sont intéressées au taux d'HbA1c chez les personnes non-diabétiques. Identifier et connaître ces facteurs peuvent aider au choix des facteurs à tester chez le patient souffrant de DT1.

Différentes maladies touchant le cycle de l'hémoglobine peuvent influencer le taux d'HbA1c, notamment celles perturbant l'érythropoïèse ou provoquant la destruction des globules rouges (Gallagher, Le Roith, & Bloomgarden, 2009; Goldenberg, Cheng, Punthakee, & Clement, 2011). Des formes particulières d'hémoglobine, présentes dans certaines populations, associées ou non à des maladies, peuvent être responsables d'un taux plus élevé d'HbA1c (Tsugawa, Mukamal, Davis, Taylor, & Wee, 2012; Venkataraman *et al.*, 2012). En 2015, Cavagnoli *et al* ont effectué une revue et une méta-analyse de l'impact de la carence ou de la déficience en fer, des variants génétiques de l'hémoglobine (dont l'hémoglobine S, forme de l'hémoglobine à l'origine de la drépanocytose) et de l'urémie sur le taux d'HbA1c. Cette étude n'a pas permis de montrer un effet significatif de ces facteurs sur le taux d'HbA1c (Cavagnoli, Pimentel, Freitas, Gross, & Camargo, 2015).

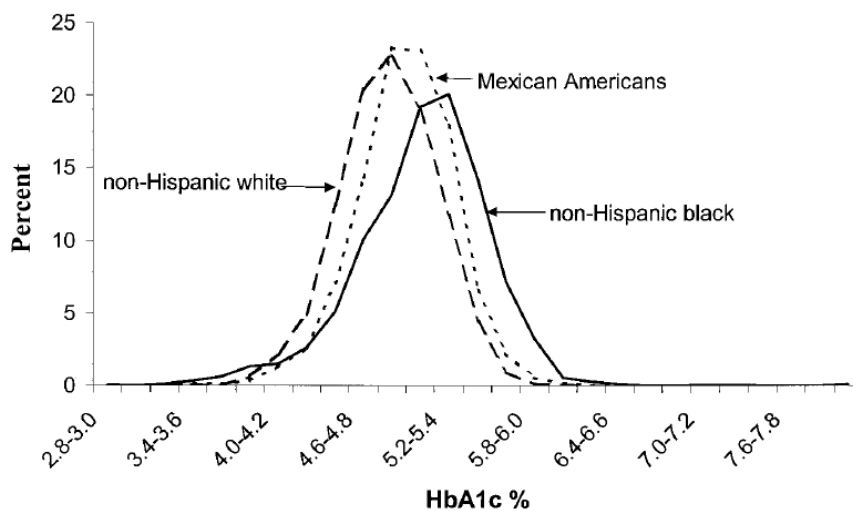
L'âge et le genre sont corrélés au taux d'HbA1c chez des individus non-diabétiques (Figure 11), les hommes ayant un taux d'HbA1c plus élevé que les femmes, particulièrement entre 30 et 45 ans (Bae *et al.*, 2014; Balkau *et al.*, 2011).

L'étude d'une cohorte américaine a montré que l'origine ethnique influence le taux d'HbA1c de façon significative, même après prise en compte de l'âge et du genre (Figure 12) (Saaddine *et al.*, 2002). Dans cette cohorte, l'âge, le genre, l'ethnie, et la présence d'histoire de diabète dans la famille ont un effet significatif sur le taux d'HbA1c et expliquent au total 6% de la variation du taux d'HbA1c.



**Figure 11 :** Variation du taux d'HbA1c en fonction de l'âge et du genre

A- Variation du taux d'HbA1c et du glucose à jeun en fonction de l'âge et du genre dans une population européenne d'après Balkau *et al* 2011 (Balkau *et al.*, 2011). FPG : *Fasting plasma glucose* ou glycémie à jeun ; B-Variation du taux d'HbA1c en fonction de l'âge et du genre dans une population coréenne d'après Bae *et al* 2014 (Bae *et al.*, 2014).

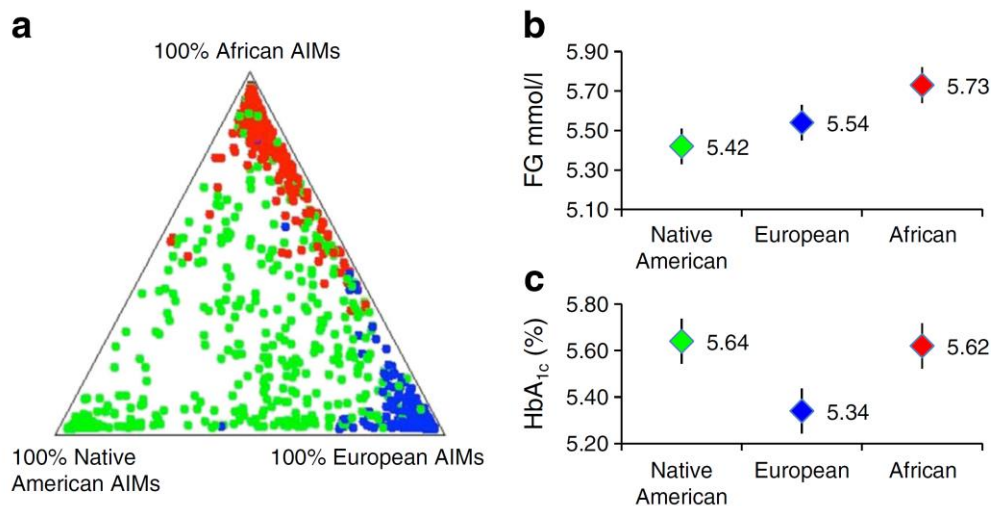


**Figure 12 :** Distribution du taux d'HbA1c en fonction de l'ethnie

Taux d'HbA1c par ethnie chez des enfants et jeunes adultes vivant aux Etats-Unis âgés entre 5 et 24 ans de la cohorte NHANES III (National Health and Nutrition Examination Survey) (1988–1994) ; d'après Saaddine *et al* 2002 (Saaddine *et al.*, 2002). *Non-Hispanic white* : Blanc non-latino ; *Mexican Americans* : Américains d'origine mexicaine ; *non-Hispanic black* : Noir non-latino

La présence de certains marqueurs génétiques, appelé « marqueurs d'origine », permet d'avoir l'information de l'origine du patient, et permet notamment de différencier les patients d'origine africaine ou européenne. La présence d'allèles associés à une origine africaine est associée à une augmentation du taux d'HbA1c d'environ 0.27% par rapport à ceux associés à une origine européenne, même après prise en compte d'autres facteurs (l'âge, le genre, l'Indice de Masse Corporel (IMC) et les marqueurs de statut socioéconomiques (SES) classiques) (Meigs *et al.*, 2014) (Figure 13). La part de variance du taux d'HbA1c expliquée par l'origine ethnique est de

l'ordre de 0.6% (Maruthur *et al.*, 2011). Cependant, l'effet de l'ethnie ou de l'origine des patients sur le taux d'HbA1c est complexe, puisque ce facteur est corrélé avec un certain nombre d'autres facteurs tels que l'IMC ou le SES (Herman, 2016; Selvin, 2016).



**Figure 13 :** Différence du taux d'HbA1c entre individus non-diabétiques d'origines différentes vivant aux Etats-Unis, d'après Meigs *et al* 2014 (Meigs *et al.*, 2014)

AIMs : *Ancestry informative markers* ou marqueurs d'origine ; FG : *fasting glucose* ou glycémie ; *African* : Africain ; *European* : Européen ; *Native American* : Amérindien

D'autres facteurs, liés au mode de vie, tel que la consommation tabac vie entière, ont été identifiés dans une méta-analyse comme influençant le taux d'HbA1c (Soulimane *et al.*, 2014).

### 1.3.4 Les facteurs génétiques du taux d'HbA1c chez les personnes non-diabétiques

En utilisant les méthodes développées dans la première et deuxième partie de ce chapitre, différentes études ont pu identifier la part génétique du taux d'HbA1c, ainsi que des marqueurs, gènes ou régions associés à une variation du taux d'HbA1c chez les individus non-diabétiques.

#### 1.3.4.1 L'héritabilité du taux d'HbA1c

L'héritabilité du taux d'HbA1c a été estimée dans différentes cohortes d'individus non-diabétiques (Tableau 5) entre 47% et 72%. Cette grande variabilité est due à la forte hétérogénéité des cohortes (origine, taille de l'échantillon, cohorte de jumeaux, familles ou cohortes de sujets).

Référence	Origine ou pays	Taille de la population	Héritabilités
Meigs <i>et al</i> 2002 (Meigs, Panhuysen, Myers, Wilson, & Cupples, 2002)	Royaume-Uni	771 individus	55% [37%-73%]
Cohen <i>et al</i> 2006 (Cohen <i>et al.</i> , 2006)	Royaume-Uni	40 MZ + 46 DZ	71% [56%-81%]
Pilia <i>et al</i> 2006 (Pilia <i>et al.</i> , 2006)	Sardaigne	6 148 individus (711 familles)	59% [54%-63%]
Simonis-Bik <i>et al</i> 2008 (Simonis-Bik <i>et al.</i> , 2008)	Caucasiens	51 MZ + 60 germains non MZ	75% [62%-84%]
Rahman <i>et al</i> 2009 (Rahman <i>et al.</i> , 2009)	Suède	1 262 MZ + 320 DZ	71%
Tejero <i>et al</i> 2010 (Tejero <i>et al.</i> , 2010)	Eskimos	1 214 individus	47% [29%-65%]
Sung <i>et al</i> 2015 (Sung, Lee, Song, Lee, & Kim, 2015)	Corée	705 MZ + 164 DZ	72% [55%-76%]

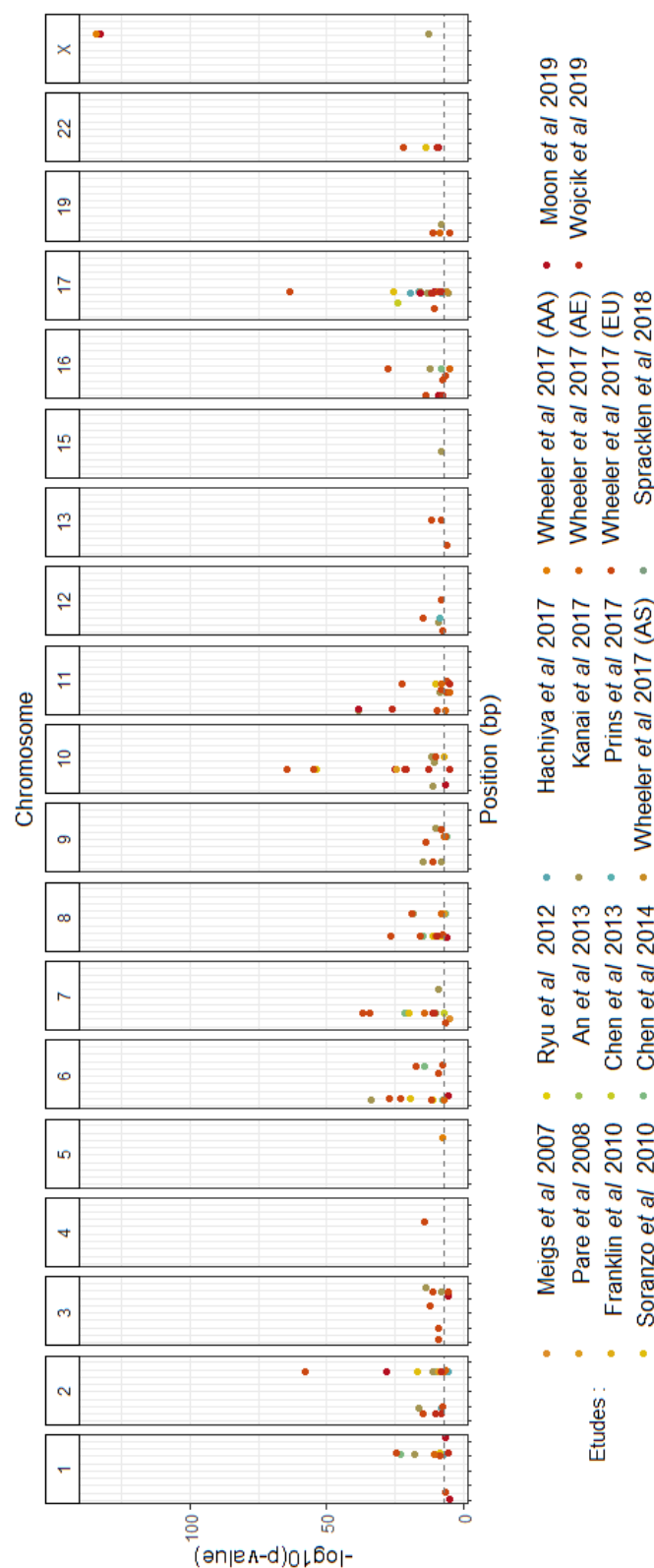
**Tableau 5** : Estimation de l'héritabilité du taux d'HbA1c chez des patients non-diabétiques  
 MZ : Jumeaux monozygotes – DZ : jumeaux dizygotes – Germain : frères et/ou sœurs

#### 1.3.4.2 Les régions et gènes influençant le taux d'HbA1c

Entre 2007 et 2015, plusieurs études ont permis d'identifier 18 *loci* associés au taux d'HbA1c chez des individus non-diabétiques (An *et al.*, 2014; Chen *et al.*, 2013; Chen *et al.*, 2014; Franklin *et al.*, 2010; Meigs *et al.*, 2007; Pare *et al.*, 2008; Ryu & Lee, 2012; Soranzo *et al.*, 2010). Depuis 2017, sept nouveaux GWAS et méta-analyses ont répliqué ces régions et ont mis en évidence plus de soixante nouvelles régions (Kanai *et al.*, 2018; Moon *et al.*, 2019; Prins *et al.*, 2017; Spracklen *et al.*, 2018; Wheeler *et al.*, 2017; Wojcik *et al.*, 2019). Les caractéristiques de ces études sont présentées dans le Tableau 6, et les résultats de ces études sont présentés dans la Figure 14. Au total, 90 gènes ou régions ont ainsi été identifiés.

A l'exception des deux études publiées en 2019, qui se concentrent sur des minorités ethniques vivant aux Etats-Unis, l'ensemble de ces études a été effectué sur (Tableau 6) :

- Des populations asiatiques (Chen *et al.*, 2013; Chen *et al.*, 2014; Kanai *et al.*, 2018; Ryu & Lee, 2012; Spracklen *et al.*, 2018) ;
- Ou européennes (An *et al.*, 2014; Franklin *et al.*, 2010; Meigs *et al.*, 2007; Pare *et al.*, 2008; Prins *et al.*, 2017; Soranzo *et al.*, 2010).



**Figure 14 :** Marqueurs associés au taux d'HbA1c dans les différentes études disponibles sur GWAS Catalog. Chaque point correspond au seuil de significativité de l'association d'un marqueur identifié dans l'une des 13 études référencées dans GWAS Catalog (MacArthur et al., 2017). Chaque couleur correspond à l'une de ces études. Extraction faite le 10 janvier 2020. La ligne en pointillé correspond au seuil classique de significativité de 5E-8.

Article	Type d'étude	Pays	Taille de la population	Nombre de <i>loci</i> identifiés*
Meigs et al. 2007 (Meigs <i>et al.</i> , 2007)	Cohorte	Etats-Unis	623	1
Pare et al. 2008 (Pare <i>et al.</i> , 2008)	Cohorte	Etats-Unis	14 618	4
Franklin et al. 2010 (Franklin <i>et al.</i> , 2010)	Méta-analyse	Europe	1782	1
Soranzo et al. 2010 (Soranzo <i>et al.</i> , 2010)	Méta-analyse	Etats-Unis Europe	46 368	11
Ryu and Lee 2012 (Ryu & Lee, 2012)	Cohorte	Corée	8 057	1
Chen et al. 2013 (Chen <i>et al.</i> , 2013)	Méta-analyse	Singapour	6 682	3
An et al. 2014 (An <i>et al.</i> , 2014)	Cohortes et étude prospective	Etats-Unis	5 656	1
Chen et al. 2014 (Chen <i>et al.</i> , 2014)	Méta-analyse	Asie de l'est	32 602	12
Wheeler et al. 2017 (Wheeler <i>et al.</i> , 2017)	Méta-analyse	Europe (EU) Afro-Américain (AA) Asie de l'est (AE) Asie du Sud (AS)	159 940	69
Prins et al. 2017 (Prins <i>et al.</i> , 2017)	Etude populationnelle	Royaume-Uni	9 961	2
Hachiya et al. 2017 (Hachiya <i>et al.</i> , 2017)	Cohorte	Japon	7 704	7
Spracklen et al. 2018 (Spracklen <i>et al.</i> , 2018)	Etude populationnelle	Chine	8 403	1
Kanai et al. 2018 (Kanai <i>et al.</i> , 2018)	Etude populationnelle	Japon	162 255	28
Wojcik et al. 2019 (Wojcik <i>et al.</i> , 2019)	Etude populationnelle	Minorité ethnique vivant aux Etats-Unis	11 178	17
Moon et al. 2019 (Moon <i>et al.</i> , 2019)	Cohorte	Hispanique et latinos vivant aux Etats-Unis	4 729	21

**Tableau 6** : GWAS et méta-analyses étudiant le taux d'HbA1c disponibles sur GWAS Catalog

\* Nombre de marqueurs reportés dans GWAS Catalog le 10 janvier 2020 au seuil de 8E-6

Deux études, publiées en 2019, se sont intéressées à la génétique de différents facteurs, dont le taux d'HbA1c, dans le contexte d'appartenance à une population minoritaire (Wojcik *et al.*, 2019) ou plus précisément à la population hispanique (Moon *et al.*, 2019) aux Etats-Unis.

La méta-analyse de Wheeler *et al* de 2017 est multi-ethnique avec des populations européennes, asiatiques et afro-africaines (Wheeler *et al.*, 2017). Les auteurs ont étudié 84 cohortes, dont 19 asiatiques (13 de l'Est et 6 du Sud de l'Asie), 9 d'origine afro-américaine et le reste d'origine européenne, en prenant en compte l'origine de la population pour effectuer la méta-analyse. Ils ont ainsi pu mettre en évidence l'association de 69 *loci* associés au taux



d'HbA1c chez des individus non-diabétiques. Les auteurs ont classé chacun de ces gènes dans l'une des trois catégories suivantes, en fonction des études d'association déjà publiées :

- Les régions associées avec au moins deux traits glycémiques ont été classées comme étant glycémiques ;
- Les régions associées avec au moins deux traits érythrocytaires ont été classées comme étant érythrocytaires ;
- Les autres régions n'étant associées à aucune de ces deux catégories ont été classées dans une catégorie notée autre.

À partir de la classification proposée dans la méta-analyse de Wheeler *et al.*, l'ensemble des marqueurs présents dans la base de données GWAS Catalog, correspondant aux quinze études précédemment présentées, a été reclassé. Vingt-cinq régions ont une fonction « glycémique » et 35 une fonction « érythrocytaire ». Les 30 autres régions n'ont soit pas de fonction connue soit une fonction n'entrant pas dans ces deux catégories.

Nous avons décidé de ne présenter que les régions qui avait été identifiées dans au moins deux études ainsi que les régions n'ayant pas un effet trop faible sur le taux d'HbA1c. Pour cela, nous avons choisi de ne présenter que les marqueurs associés au taux d'HbA1c avec un coefficient de régression estimé, sous un mode de transmission additif soit supérieur ou égal en valeur absolue à 0.05 ou à 0.01 sur au moins deux études indépendantes. Parmi les 90 régions, nous avons ainsi sélectionné 41 régions. Ces différentes régions sont présentées dans le Tableau 7 pour les régions glycémiques, le Tableau 8 pour les régions érythrocytaires et pour les autres régions dans le Tableau 9.

Deux autres études ont étudié les facteurs génétiques du taux d'HbA1c dans des populations particulières : une population arabe (Hebbar *et al.*, 2017) et une population Inuit (Appel *et al.*, 2018).

Dans les populations africaines et Inuits, les effets de certains *loci* semblent différents (Appel *et al.*, 2018; Wheeler *et al.*, 2017). Pour le marqueur rs6474359 du gène ANK1, l'allèle C est associé à une diminution du taux d'HbA1c dans les populations européennes, alors qu'il est associé avec un taux d'HbA1c plus élevé chez les Inuits (Appel *et al.*, 2018). Dans le cas des populations africaines, la majorité de la variation du taux d'HbA1c peut s'expliquer par un seul variant

présent sur le chromosome X dans le gène G6PD, contrairement aux populations européennes et asiatiques (Wheeler *et al.*, 2017). Ces deux résultats renforcent l'importance d'étudier des populations relativement uniformes et de prendre en compte l'ethnie dans les analyses génétiques.

Les cliniciens considèrent que le taux d'HbA1c a significativement évolué chez le sujet souffrant de diabète quand celui-ci a varié d'au moins 0.5%. Or, ici, aucune région ne semble avoir cet impact sur le taux d'HbA1c. Quatre régions (G6PD, ASB16, YKT6-GCK et SOX30) ont un impact supérieur à 0.1%, dont deux uniquement dans les populations afro-américaines (G6PD et SOX30).

Chr	Gène d'intérêt	Position (bp)	Variant et allèle à risque	$\beta$ [Intervalle de confiance]	Origine	Pvalue	Référence	
2	G6PC2, SPC25	168898336	rs13387347-T	-0.020 [0.013-0.028]	Asie	1.0E-07	(Wheeler <i>et al.</i> , 2017)	
		168900844	rs1402837-A	0.02	Europe	5.0E-10	(Pare <i>et al.</i> , 2008)	
	G6PC2	168906638	rs560887-C	0.028 [0.025-0.032]	Europe	1.0E-58	(Wheeler <i>et al.</i> , 2017)	
				-0.032 [0.024-0.040]	Europe	1.0E-17	(An <i>et al.</i> , 2014)	
	ABCB11, G6PC2	168910638	rs12053049		-0.11 [0.074-0.148]	Minorité	5.0E-9	(Wojcik <i>et al.</i> , 2019)
					0.048 [0.034-0.061]	Asie	1.0E-11	(Kanai <i>et al.</i> , 2018)
				0.03 [0.01-0.05]	Hispanique	6.0E-29	(Moon <i>et al.</i> , 2019)	
				0.05 [0.04-0.06]	Europe	8.0E-18	(Soranzo <i>et al.</i> , 2010)	
ABCB11	168934928	rs557462-C	0.0274 [0.014-0.033]	Asie	4.0E-06	(Hachiya <i>et al.</i> , 2017)		
	168935661	rs3755157-T	0.07 [0.050-0.090]	Asie	3.0E-11	(Chen <i>et al.</i> , 2014)		
3	SLC2A2	171007094	rs8192675-T	0.011 [0.0079-0.0145]	Europe	1.0E-11	(Wheeler <i>et al.</i> , 2017)	
		171014810	rs7356034	-0.051 [0.034-0.067]	Asie	5.0E-09	(Kanai <i>et al.</i> , 2018)	
	IGF2BP2	185801133	rs76922886	0.057 [0.042-0.071]	Asie	2.0E-14	(Kanai <i>et al.</i> , 2018)	
		20665715	rs7772603-C	0.06 [0.04-0.08]	Asie	4.0E-08	(Chen <i>et al.</i> , 2014)	
6	CDKAL1	20675561	rs35261542	0.087 [0.073-0.101]	Asie	2.0E-34	(Kanai <i>et al.</i> , 2018)	
		20679478	rs7756992-G	0.012 [0.009-0.016]	Europe	3.0E-12	(Wheeler <i>et al.</i> , 2017)	
	GCK	20725192	rs7747752-C	0.019 [0.012-0.026]	Asie	1.0E-07	(Wheeler <i>et al.</i> , 2017)	
		44184122	rs730497	-0.046	Asie	1.0E-11	(Ryu & Lee, 2012)	
7	GCK	44186502	rs2971670-T	0.11 [0.082-0.147]	Minorité	5.0E-12	(Wojcik <i>et al.</i> , 2019)	
		44189469	rs1799884	0.04 [0.02-0.06]	Hispanique	3.0E-11	(Moon <i>et al.</i> , 2019)	
	YKT6, GCK	44189469	rs1799884-T	[0.079-0.14]	Minorité	2.0E-11	(Wojcik <i>et al.</i> , 2019)	
				0.04 [0.03-0.05]	Europe	1.0E-20	(Soranzo <i>et al.</i> , 2010)	
		44192287	rs6975024	0.12 [0.10-0.14]	Asie	2.0E-22	(Chen <i>et al.</i> , 2014)	
		44196069	rs4607517-A	0.44 [0.42-0.46]	Asie	6.0E-08	(Chen <i>et al.</i> , 2014)	
				0.059 [0.042-0.077]	Asie	4.0E-11	(Kanai <i>et al.</i> , 2018)	
				0.031 [0.026-0.035]	Europe	9.0E-38	(Wheeler <i>et al.</i> , 2017)	



Chr	Gène d'intérêt	Position (bp)	Variant et allèle à risque	$\beta$ [Intervalle de confiance]	Origine	Pvalue	Référence
1	SMG5	156263398	rs2273833	-0.074 [0.058-0.091]	Asie	1.0E-18	(Kanai <i>et al.</i> , 2018)
		156276128	rs144991356-CT	0.033 [0.021-0.045]	Asie	3.0E-08	(Hachiya <i>et al.</i> , 2017)
	TMEM79, SMG5	156285665	rs6684514-G	0.031 [0.02-0.043]	Asie	1.0E-07	(Hachiya <i>et al.</i> , 2017)
				0.09 [0.07-0.11]	Asie	1.0E-23	(Chen <i>et al.</i> , 2014)
		158615625	rs2779116-T	0.02 [0.01-0.03]	Europe	3.0E-09	(Soranzo <i>et al.</i> , 2010)
2	KLHL29	158656588	rs857691-T	0.067 [0.039-0.096]	Minorité	3.0E-6	(Wojcik <i>et al.</i> , 2019)
		23740319	rs193292177	0.019 [0.016-0.023]	Europe	4.0E-25	(Wheeler <i>et al.</i> , 2017)
		26092913	rs1800562-G	0.39 [0.27-0.5]	Minorité	4.0E-11	(Wojcik <i>et al.</i> , 2019)
				0.040 [0.033-0.047]	Europe	5.0E-28	(Wheeler <i>et al.</i> , 2017)
				0.06 [0.05-0.07]	Europe	3.0E-20	(Soranzo <i>et al.</i> , 2010)
6	HFE	135097778	rs7776054	-0.063 [0.049-0.077]	Asie	3.0E-18	(Kanai <i>et al.</i> , 2018)
		135097880	rs9399137-T	0.07 [0.050-0.090]	Asie	9.0E-15	(Chen <i>et al.</i> , 2014)
		135110502	rs11154792-T	0.0205 [0.016-0.025]	Europe	7.00E-18	(Wheeler <i>et al.</i> , 2017)
		41691676	rs6474359-T	0.044 [0.033-0.054]	Europe	2.0E-16	(Wheeler <i>et al.</i> , 2017)
		41772887	rs4737009-A	0.06 [0.04-0.08]	Europe	1.0E-08	(Soranzo <i>et al.</i> , 2010)
8	ANK1			0.021 [0.017-0.025]	Europe	4.0E-27	(Wheeler <i>et al.</i> , 2017)
				0.026 [0.018-0.033]	Asie	1.0E-10	(Wheeler <i>et al.</i> , 2017)
				0.03 [0.02-0.04]	Europe	6.0E-12	(Soranzo <i>et al.</i> , 2010)
				0.08 [0.060-0.100]	Asie	1.0E-15	(Chen <i>et al.</i> , 2014)
				0.03 [0.01-0.05]	Hispanique	5.0E-10	(Moon <i>et al.</i> , 2019)
				0.093 [0.064-0.12]	Minorité	3.0E-10	(Wojcik <i>et al.</i> , 2019)
				0.095 [0.084-0.106]	Europe	3.0E-65	(Wheeler <i>et al.</i> , 2017)
				0.033 [0.028-0.037]	Europe	2.0E-55	(Wheeler <i>et al.</i> , 2017)
				0.09 [0.08-0.10]	Europe	3.0E-54	(Soranzo <i>et al.</i> , 2010)
				0.16 [0.12-0.20]	Minorité	2.0E-13	(Wojcik <i>et al.</i> , 2019)
10	HK1	69330087	rs4745982-T	0.01 [0.08-0.12]	Hispanique	6.0E-26	(Moon <i>et al.</i> , 2019)
		69331257	rs10823343-A				
		69333636	rs16926246-C				
		69339353	rs72805692-G				

						0.27 [0.21-0.32]	Minorité	1.0E-21	(Wojcik <i>et al.</i> , 2019)
		69340157	rs7072268-A			0.05	Europe	2.0E-25	(Pare <i>et al.</i> , 2008)
11	HBB	5227002	rs334-A			0.061 [0.034-0.088]	Minorité	8.0E-06	(Wojcik <i>et al.</i> , 2019)
						0.305 [0.26-0.35]	Hispanique	3.0E-39	(Moon <i>et al.</i> , 2019)
						0.82 [0.67-0.97]	Minorité	8.0E-27	(Wojcik <i>et al.</i> , 2019)
16	HBM	170584	rs145546625-T			0.15 [0.098-0.204]	Minorité	4.0E-08	(Wojcik <i>et al.</i> , 2019)
	PIEZO1	88778524	rs9933309-C			0.059 [0.039-0.079]	Hispanique	1.0E-09	(Moon <i>et al.</i> , 2019)
	LOC107984842, CDT1	88792238	rs2968478-?			0.07 [0.050-0.090]	Asie	1.0E-08	(Chen <i>et al.</i> , 2014)
						0.050 [0.036-0.064]	Asie	1.0E-12	(Kanai <i>et al.</i> , 2018)
17	TMC6	78125783	rs2748427-G			0.057 [0.045-0.07]	Asie	5.0E-20	(Hachiya <i>et al.</i> , 2017)
						0.078 [0.058-0.098]	Asie	7.0E-14	(Kanai <i>et al.</i> , 2018)
		78125997	rs761772-C			-	Europe	5.0E-09	(Prins <i>et al.</i> , 2017)
		78128784	rs2748424-G			0.051 [0.039-0.063]	Hispanique	1.0E-16	(Moon <i>et al.</i> , 2019)
		78131350	rs383603			0.11 [0.079-0.137]	Minorité	8.0E-12	(Wojcik <i>et al.</i> , 2019)
19	MYO9B	17121689	rs11667918-C			0.06 [0.040-0.080]	Asie	9.0E-12	(Chen <i>et al.</i> , 2014)
		17135927	rs11086054-A			0.0097 [0.0054-0.014]	Europe	8.0E-06	(Wheeler <i>et al.</i> , 2017)
						0.024 [0.016-0.031]	Asie	2.0E-09	(Wheeler <i>et al.</i> , 2017)
		17145713	rs17533903-A			0.015 [0.011-0.019]	Europe	5.0E-12	(Wheeler <i>et al.</i> , 2017)
22	TMPRSS6	37066896	rs855791-A			0.03 [0.02-0.04]	Europe	3.0E-14	(Soranzo <i>et al.</i> , 2010)
						0.086 [0.06-0.113]	Minorité	3.0E-10	(Wojcik <i>et al.</i> , 2019)
						0.03 [0.01-0.05]	Hispanique	9.0E-10	(Moon <i>et al.</i> , 2019)
		37073551	rs4820268-G			0.016 [0.013-0.02]	Europe	1.0E-22	(Wheeler <i>et al.</i> , 2017)
X	G6PD	154536002	rs1050828-T			-0.38 [0.35-0.41]	Afro-Américain	8.0E-135	(Wheeler <i>et al.</i> , 2017)
						0.346 [0.32-0.37]	Hispanique	1.0E-133	(Moon <i>et al.</i> , 2019)

**Tableau 8 :** Régions considérées comme étant érythrocytaires et qui ont été identifiées dans au moins deux études, ou avec un coefficient de régression supérieur à 0.05

Chr	Gène d'intérêt	Position (bp)	Variant et allèle à risque	$\beta$ [Intervalle de confiance]	Origine	Pvalue	Référence
1	TM1EM206	212371300	rs199584848-A	0.08 [0.051-0.109]	Hispanique	2.0E-07	(Moon <i>et al.</i> , 2019)
2	SIX3, SIX2	44964861	rs10168523-G	0.029 [0.019-0.039]	Asie	9.0E-09	(Hachiya <i>et al.</i> , 2017)
			rs12712928	0.060 [0.046-0.074]	Asie	4.0E-17	(Kanai <i>et al.</i> , 2018)
5	SOX30	157628483	rs11954649-G	0.12 [0.077-0.16]	Afro-Américain	4.0E-08	(Wheeler <i>et al.</i> , 2017)
6	LOC107986647, LOC105378010	135097778	rs7776054	-0.063 [0.049-0.077]	Asie	3.0E-18	(Kanai <i>et al.</i> , 2018)
		135097880	rs9399137-T	0.07 [0.050-0.090]	Asie	9.0E-15	(Chen <i>et al.</i> , 2014)
7	HLA-DQA2	32701596	rs11752643-T	0.096 [0.061-0.131]	Hispanique	2.0E-06	(Moon <i>et al.</i> , 2019)
		128201573	rs4728092	0.073 [0.049-0.096]	Asie	1.0E-09	(Kanai <i>et al.</i> , 2018)
8	UNC5D	36068884	rs6992753-A	0.111 [0.072-0.15]	Hispanique	1.0E-06	(Moon <i>et al.</i> , 2019)
9	KLF4	107774651	rs1467311-G	0.0086 [0.0053-0.012]	Europe	2.0E-07	(Wheeler <i>et al.</i> , 2017)
				0.024 [0.015-0.032]	Asie	1.0E-07	(Wheeler <i>et al.</i> , 2017)
				0.04 [0.020-0.060]	Asie	1.0E-06	(Chen <i>et al.</i> , 2014)
10	INPP5E	136434270	rs7851507	-0.051 [0.036-0.067]	Asie	8.0E-11	(Kanai <i>et al.</i> , 2018)
				0.12 [0.073-0.163]	Hispanique	2.0E-07	(Moon <i>et al.</i> , 2019)
				-0.069 [0.049-0.089]	Asie	2.0E-11	(Kanai <i>et al.</i> , 2018)
12	FAM60A	31313679	rs147538848	0.058 [0.039-0.076]	Asie	8.0E-10	(Kanai <i>et al.</i> , 2018)
				-0.36 [0.29-0.43]	Asie	1.0E-24	(Chen <i>et al.</i> , 2013)
17	ASB16, C17orf53	44164561	rs12602486-G	0.041 [0.031-0.051]	Asie	6.0E-17	(Hachiya <i>et al.</i> , 2017)
				0.08 [0.060-0.100]	Asie	4.0E-17	(Chen <i>et al.</i> , 2014)
				0.086 [0.053-0.119]	Asie	4.0E-07	(Spracklen <i>et al.</i> , 2018)
FN3KRP	82723985	82727550	rs35203608-CA	0.080 [0.053-0.107]	Minorité	4.0E-09	(Wojcik <i>et al.</i> , 2019)
				0.023 [0.015-0.03]	Asie	7.0E-10	(Wheeler <i>et al.</i> , 2017)
				0.028 [0.025-0.031]	Europe	4.0E-64	(Wheeler <i>et al.</i> , 2017)
			rs1046896-T	0.031 [0.019-0.043]	Asie	2.0E-07	(Chen <i>et al.</i> , 2013)
				0.040 [0.03-0.049]	Asie	2.0E-16	(Hachiya <i>et al.</i> , 2017)
				0.04 [0.03-0.05]	Europe	2.0E-26	(Soranzo <i>et al.</i> , 2010)

				0.052 [0.031-0.072]	Asie	7.0E-07	(Wheeler et al., 2017)
FN3K				0.080 [0.053-0.107]	Minorité	8.0E-09	(Wojcik et al., 2019)
	82735405	rs2256339-T		0.03 [0.01-0.05]	Hispanique	2.0E-11	(Moon et al., 2019)
	82736950	rs8067360		0.057 [0.044-0.071]	Asie	1.0E-16	(Kanai et al., 2018)

**Tableau 9** : Régions considérées comme étant autre qui ont été identifiées dans au moins deux études, ou avec un coefficient de régression supérieur à 0.05



# Chapitre 2 | Étude des propriétés statistiques du « General Regression Model » (GRM) dans le cas d'un trait quantitatif

---

## 2.1 Résumé

Avant mon arrivée, une étude de simulation modélisant un trait binaire avait montré que le test GRM, qui permet de tester l'association génétique sans faire d'hypothèse sur le mode de transmission, pouvait être plus puissant que les tests basés sur des modèles de régression logistique. De plus, ce travail a montré que le GRM permet de tester formellement le modèle génétique sous-jacent (Dizier *et al.*, 2017).

Dans le cadre de mon travail de thèse, nous avons effectué une étude de simulation afin de comparer les propriétés statistiques du GRM aux trois LR (LR\_ADD, LR\_DOM, et LR\_REC) dans le cas d'un trait quantitatif. Nous avons également comparé ses propriétés à un test MAX paramétrique (Loley *et al.*, 2013) et à un test allélique développé pour les traits quantitatifs (Lee *et al.*, 2013). Nous avons ensuite appliqué le GRM aux données du taux d'HbA1c de la cohorte de volontaires sains français D.E.S.I.R. afin de voir si de nouveaux *loci* pouvaient être identifiés.

Pour un trait quantitatif, le GRM est basé sur un modèle de régression linéaire, comme montré précédemment. Les génotypes sont modélisés par deux variables indicatrices Add et DomDev comme proposé par Fisher (Fisher, 1918) et Wilson (Wilson, 1980) (Tableau 10). La variable Add permet de modéliser l'effet dose de l'allèle à risque et la variable DomDev modélise l'écart à l'additivité. Le GRM permet également de prendre en compte des covariables (COV) dans l'analyse. Pour un phénotype quantitatif (P), le modèle s'écrit (avec E le terme d'erreur) :

$$P = \alpha + \beta_{Add} \cdot Add + \beta_{DomDev} \cdot DomDev + \beta_{COV} \cdot COV + E$$

Sous  $H_0$  ( $\beta_{Add} = \beta_{DomDev} = 0$ ), le rapport de vraisemblance entre le modèle avec uniquement les covariables et le modèle avec l'ajout des deux variables Add et DomDev suit une loi de  $\chi^2$  à 2 degrés de liberté. Le mode de transmission génétique est ensuite testé pour les SNPs significativement associés ( $p \leq$  seuil de significativité calculé par rapport au nombre de tests effectués).

Génotypes	Add	DomDev
AA	0	0
Aa	1	1
Aa	2	0

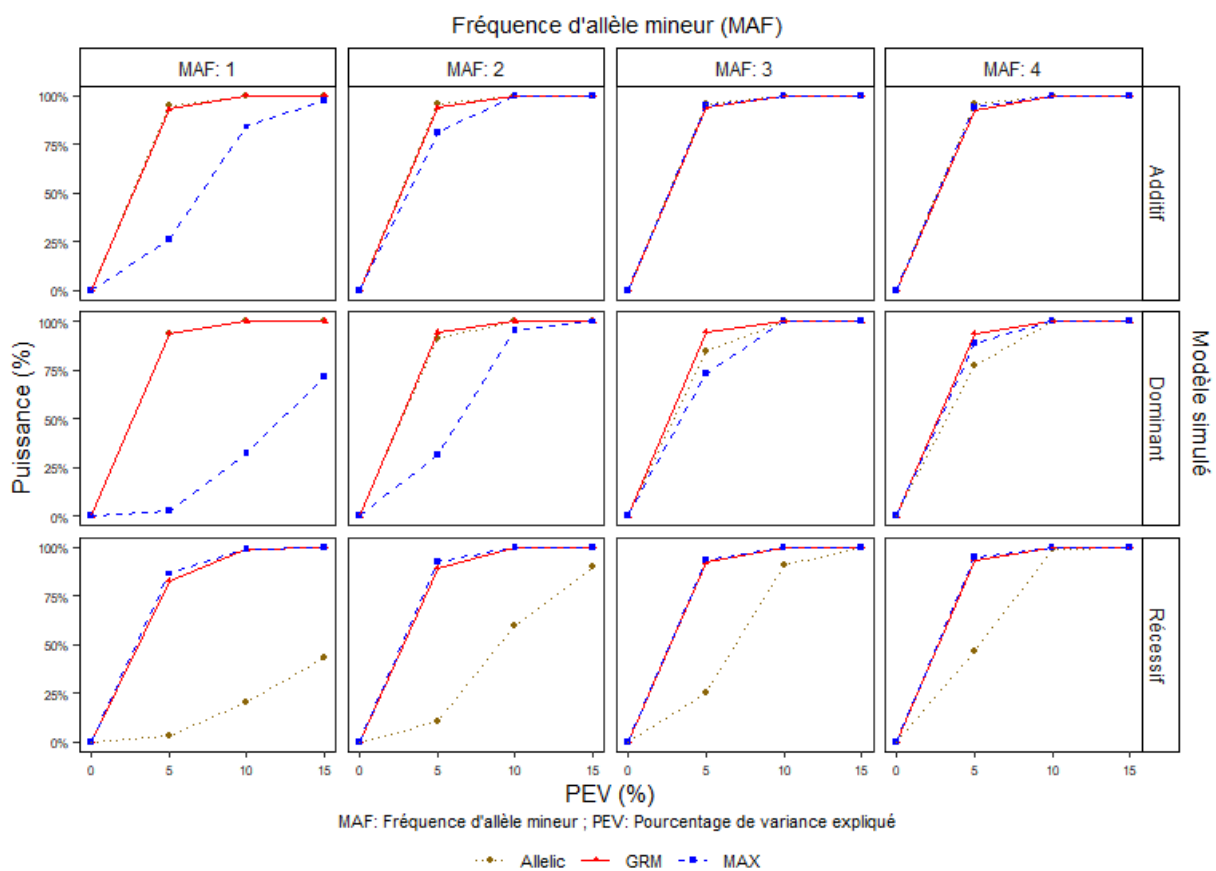
**Tableau 10** : Codage des variables en fonction du génotype

Dans un premier temps, le modèle additif est testé, sous  $H_0$ ,  $\beta_{DomDev}=0$ . Si  $\beta_{DomDev}$  est significativement différent de 0, les modèles dominant, récessif et d'*over-dominance* seront alors testés. Si  $\beta_{DomDev}$  appartient à l'intervalle de confiance de  $\beta_{Add}$  ( $\beta_{DomDev} \in [\beta_{Add} \pm Z_{\alpha/2} \times \sigma_{Add}]$ ), le modèle dominant sera retenu. Si  $\beta_{DomDev}$  appartient à l'intervalle de confiance de  $-\beta_{Add}$  ( $\beta_{DomDev} \in [-\beta_{Add} \pm Z_{\alpha/2} \times \sigma_{Add}]$ ), le modèle récessif sera retenu. Il est important de remarquer que plus le risque de première espèce  $\alpha$  est petit plus l'intervalle de confiance est grand. Si  $\beta_{DomDev}$  est significativement différent de 0 et  $\beta_{Add}$  n'est pas significativement différent de 0, alors le modèle d'*over-dominance* sera retenu. L'arbre de décision est présenté dans la Figure 1 de l'article.

Nous avons simulé un trait quantitatif sous différents modèles génétiques (additif, dominant et récessif), en faisant varier la fréquence de l'allèle mineur (MAF de 0.1 à 0.4) et le pourcentage de variance expliquée (PEV de 0 à 20%) pour un échantillon de 1000 individus. Nous avons analysé ces simulations avec le GRM et cinq différents tests : LR\_ADD, LR\_REC, et LR\_DOM, le Test Allélique quantitatif (*Allelic*) (Lee *et al.*, 2013) et une version du test MAX (Loley *et al.*, 2013).

La robustesse du GRM est comparable à celles des trois LR (Table 1, supplementary data de l'article en Annexe 1), et concernant la puissance, le GRM est au moins aussi puissant que le LR\_ADD pour l'ensemble des modèles testés (Figure 2 de l'article). En effet, sous un modèle additif, l'écart de puissance entre le LR\_ADD et le GRM est au maximum de 8.6%, le LR\_ADD étant autant voire plus puissant ; alors que sous un modèle dominant ou récessif le GRM est

toujours plus puissant, avec un écart maximum de 24.1% pour le modèle dominant pour des MAF importantes et de 80% pour un modèle récessif pour de petites MAFs. L'écart de puissance entre la régression sous le modèle simulé et le GRM est toujours inférieur à 9% (entre le GRM et le LR\_DOM sous un modèle dominant et entre le GRM et le LR\_REC sous un modèle récessif). La comparaison entre le GRM et le test MAX sur un sous-échantillon de 1000 répliques, montre que le GRM est plus puissant que le MAX, pour tous les modèles simulés, notamment dans le cas d'un modèle dominant avec une différence de puissance atteignant 91% dans le cas d'une MAF de 0,1 et une PEV de 5% (Figure 15).



**Figure 15 :** Puissance du GRM et des tests MAX (Loley *et al.*, 2013) et allélique (Allelic) (Lee *et al.*, 2013) sur 1000 simulations de 1000 individus pour un seuil de  $5E-8$  en fonction de la MAF et de la PEV

La comparaison entre le test allélique et le GRM montre que le GRM est aussi voire plus puissant que le test allélique dans le cadre d'un modèle additif et dominant pour toutes les MAFs et tous les PEVs testés (Figure 15). Dans le cas du modèle récessif, le GRM est beaucoup plus puissant, notamment dans le cas d'une MAF de 0,1 et d'un PEV de 5%, avec une différence de puissance

de 79,5%. Concernant le temps d'exécution, le GRM et le test allélique sont équivalents (quelques secondes pour 1000 répliques).

En ce qui concerne le modèle génétique retenu par le GRM, dans le cas des traits simulés sous un modèle additif ou récessif, le GRM conclue au mode de transmission simulé dans la majorité des cas : pour au moins 98.4% des marqueurs simulés pour un modèle additif et pour plus de 88.9% des marqueurs pour un modèle récessif. Pour un modèle simulé dominant, et une MAF de 0.4, le mode de transmission retenu est bien dominant pour plus de 64% des marqueurs significativement associés (Figure 3 de l'article). Pour des MAF plus petites, le mode de transmission correspond au modèle simulé pour des PEVs supérieurs à 10%. Pour des PEV plus faibles, le modèle retenu est plus généralement le modèle additif.

La comparaison du GRM avec le test MAX, en dehors du modèle simulé dominant, montre que le test MAX permet de retenir moins souvent le mode de transmission simulé que le GRM, et en particulier dans le cas d'un mode de transmission simulé additif, où le mode de transmission majoritairement retenu est dominant (Figure 4 de l'article). Le GRM a un temps d'exécution identique à celui des différents tests LR (quelques secondes), alors que le temps d'exécution du test MAX est 22 fois plus longue.

La comparaison de puissance du GRM et du LR\_ADD nous a permis de confirmer le gain de puissance du GRM pour un trait quantitatif dans le cas d'écart à l'additivité, ainsi qu'une puissance comparable dans le cas d'un modèle additif. Ces résultats sont comparables aux résultats qui ont été observés pour le GRM dans le cas d'un trait binaire (Dizier *et al.*, 2017). Le GRM est aussi ou plus puissant que le test MAX, pour un modèle dominant et que le test allélique pour un modèle récessif. De plus, le GRM permet de tester formellement le mode de transmission génétique sous-jacent.

Le test GRM possède aussi des limites. Le GRM a été développé pour l'étude d'association pangénomique, c'est-à-dire pour l'analyse de SNPs fréquents ( $MAF > 0,01$ ) dans la population. Dans le cas de MAF ou de PEV trop faibles, les conditions d'application de la régression ne sont pas toujours satisfaites. Dans ce cas, il est préférable d'utiliser d'autres types de modèles, notamment les modèles non-paramétriques (Jonckheere, 1954; Terpstra, 1952), malgré le fait

qu'ils nécessitent un temps de calcul très long, tel que le test NMAX qui nécessite environ une heure pour l'analyse de 1000 SNPs sur un échantillon de 1000 personnes.

Nous avons ensuite appliqué le GRM et le LR\_ADD aux données d'HbA1c d'un sous-groupe de 695 patients normo-glycémiques de la cohorte D.E.S.I.R. (Balkau, 1996; Eschwege & Balkau, 2003), dans le cadre d'une collaboration avec l'UMR 8199 (Génomique Intégrative et Modélisation des Maladies Métaboliques) et le groupe D.E.S.I.R.. Nous avons ainsi pu suggérer un *locus* associé au taux d'HbA1c supplémentaire avec le GRM par rapport au LR\_ADD. En effet, le marqueur rs6468861 (pour lequel le GRM retient un mode de transmission additif) est aussi retenu par le LR\_ADD classique, alors que le second marqueur, le rs4594032, pour lequel le mode de transmission retenu correspond à un mode de transmission *over-dominant*, n'est pas retenu avec le LR\_ADD (Table 2 et Table 3 de l'article). Ces deux marqueurs n'avaient jamais été identifiés dans des précédentes études. Le rs6468861 est situé dans une région qui contrôle l'expression de l'hémoglobine fœtale (Garner, Tatu, Best, Creary, & Thein, 2002) qui pourrait avoir un rôle dans la formation de l'hémoglobine et donc de l'HbA1c. Le second marqueur, pour lequel le modèle retenu est un modèle d'avantage aux hétérozygotes, a été décrit dans l'étude *DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium* comme étant une région de susceptibilité au diabète de type 2 (Mahajan *et al.*, 2014) et se trouve en déséquilibre de liaison avec une région qui contient plusieurs gènes, dont le gène *MNTR1B (Melatonin receptor 1B)*, connu pour être associé à plusieurs traits glycémiques (Bonfond *et al.*, 2012; Bouatia-Naji *et al.*, 2009; Dupuis *et al.*, 2010; Lyssenko *et al.*, 2009; Prokopenko *et al.*, 2009). L'application du GRM sur les données de la cohorte D.E.S.I.R. nous a permis d'illustrer le gain de puissance du GRM pour détecter l'association par rapport au test LR\_ADD, notamment dans le cas d'écart à l'additivité.

## 2.2 Article 1 (Publié en ligne en décembre 2019)




Received: 20 July 2018 | Revised: 19 November 2019 | Accepted: 20 November 2019

DOI: 10.1111/ahg.12372

### ORIGINAL ARTICLE

Annals of  
human genetics WILEY

# General regression model: A “model-free” association test for quantitative traits allowing to test for the underlying genetic model

Emilie Gloaguen<sup>1,2</sup>  | Marie-Hélène Dizier<sup>3,4</sup> | Mathilde Boissel<sup>5,6,7</sup>  |  
Ghislain Rocheleau<sup>5,6,7</sup> | Mickaël Canouil<sup>5,6,7</sup> | Philippe Froguel<sup>5,6,7,8</sup> | Jean Tichet<sup>9</sup> |  
Ronan Roussel<sup>10,11,12</sup> | The D.E.S.I.R. Study Group<sup>1\*</sup> | Cécile Julier<sup>1,2†</sup> |  
Beverley Balkau<sup>13†</sup> | Flavie Mathieu<sup>14,15</sup> 

<sup>1</sup>Inserm UMRS-958, Paris, France

<sup>2</sup>Université Paris Diderot, Sorbonne Paris Cité, Paris, France

<sup>3</sup>Inserm UMR-946, Paris, France

<sup>4</sup>Université Paris Diderot, Sorbonne Paris Cité, Paris, France

<sup>5</sup>Université de Lille, UMR 8199 - EGID, Lille, France

<sup>6</sup>CNRS, Paris, France

<sup>7</sup>Institut Pasteur de Lille, Lille, France

<sup>8</sup>Department of Genomics of Common Disease, Imperial College London, London, United Kingdom

<sup>9</sup>IRSA, La Riche, France

<sup>10</sup>Inserm U1138, Centre de Recherche des Cordeliers, Paris, France

<sup>11</sup>Université Paris Diderot, Sorbonne Paris Cité, Paris, France

<sup>12</sup>Diabetology, Endocrinology and Nutrition Department, DHU FIRE, Hôpital Bichat, AP-HP, Paris, France

<sup>13</sup>Inserm U-1018, CESP, Team 5, UVSQ-UP, Villejuif, France

<sup>14</sup>Mission Associations Recherche & Société - Inserm Siège, DISC, Paris, France

<sup>15</sup>Paris Diderot, Sorbonne Paris Cité, Paris, France

### Correspondence

Flavie Mathieu, Inserm Siège – DISC, Mission Associations Recherche & Société, 101 Rue de Tolbiac, 75013 Paris, France.  
Email: flavie.mathieu@inserm.fr

### Funding information

This work was supported by the Institut National pour la Recherche Médicale (Inserm) and E.G. is supported by the French Ministère de la Recherche (Doctoral fellowship 2014–2017). The D.E.S.I.R. study has been supported by Inserm contracts with CNAMTS, Lilly, Novartis Pharma, and Sanofi-Aventis, by Inserm (Réseaux en Santé Publique, Interactions entre les déterminants de la santé),

### Abstract

Most genome-wide association studies used genetic-model-based tests assuming an additive mode of inheritance, leading to underpowered association tests in case of departure from additivity. The general regression model (GRM) association test proposed by Fisher and Wilson in 1980 makes no assumption on the genetic model. Interestingly, it also allows formal testing of the underlying genetic model. We conducted a simulation study of quantitative traits to compare the power of the GRM test to the classical linear regression tests, the maximum of the three statistics (MAX), and the allele-based (allelic) tests. Simulations were performed on two samples sizes, using a large panel of genetic models, varying genetic models, minor allele frequencies, and the percentage of explained variance. In case of departure from additivity, the GRM

Cohortes Santé TGIR, Association Diabète Risque Vasculaire, Fédération Française de Cardiologie, Fondation de France, ALFE-DIAM, CERINCNIEL, ONIVINS, Ardix Medical, Bayer Diagnostics, Becton Dickinson, Cardionics, Merck Santé, Novo Nordisk, Pierre Fabre, Roche, and Topcon.

\* A complete list of D.E.S.I.R. Study Group members can be found below.

† These authors contributed equally to this work.

was more powerful than the additive regression tests (power gain reaching 80%) and had similar power when the true model is additive. GRM was also as or more powerful than the MAX or allelic tests. The true simulated model was mostly retained by the GRM test. Application of GRM to HbA1c illustrates its gain in power. To conclude, GRM increases power to detect association for quantitative traits, allows determining the genetic model and is easily applicable.

#### KEYWORDS

genetic model, genome-wide association studies (GWAS), HbA1c, linear regression, quantitative traits

## 1 | INTRODUCTION

Most genome-wide association studies (GWAS) conducted in the last decades on complex diseases (McCarthy et al., 2008; Visscher, Brown, McCarthy, & Yang, 2012; Welter et al., 2014) assumed an additive genetic mode of transmission resulting in a significant loss of power when there is a strong departure from additivity (Freidlin, Zheng, Li, & Gastwirth, 2002; Schaid, McDonnell, Hebbiring, Cunningham, & Thibodeau, 2005; Slager & Schaid, 2001).

We recently proposed to use the assumption-free general regression model (GRM) association test for the case-control design, showing that GRM was more powerful than the Cochran–Armitage Trend Additive Test when there is a departure from additivity (Dizier, Demenais, & Mathieu, 2017). For quantitative traits, association is usually tested by using genetic-model-based linear regression (LR) models (Balding, 2006; Lettre, Lange, & Hirschhorn, 2007). However, LR, assuming an additive model (LR\_ADD) test, loses power to detect association in the case of deviation from additivity for binary (Freidlin et al., 2002; Schaid et al., 2005; Slager & Schaid, 2001) and quantitative traits (Lettre et al., 2007). Genetic-model-free approaches as the allele-based test have recently been developed for quantitative traits (Lee, Karrison, Cox, & Im, 2013).

Other approaches have been proposed, testing simultaneously the three different modes of inheritance (additive, dominant, and recessive) (Kitsche, Ritz, & Hothorn, 2016; Loley, König, Hothorn, & Ziegler, 2013; So & Sham, 2011), showing relatively high power but are computationally quite expensive. Recently, Loley et al. proposed a new version of one of these tests, the MAX test, based on the maximum of the three statistics (under the recessive, additive, and dominant models) (Kitsche et al., 2016; Loley et al., 2013; So & Sham, 2011). This MAX test used the generalized linear model and was thus robust for association testing and genetic model selection (Kitsche et al., 2016; Loley et al., 2013; So & Sham, 2011).

However, although several papers described well the loss of power to detect association by using the additive model in case of departure from additivity for binary traits (Dizier

et al., 2017; Freidlin et al., 2002; Schaid et al., 2005; Slager & Schaid, 2001), only one has been performed for quantitative traits (Lettre et al., 2007) considering a limited number of genetic models in their simulation study.

Our aim was first to conduct a power comparison to detect association between the GRM and the LR for a quantitative trait, under various genetic models, assuming a causal variant of the disease under additive, dominant, or recessive mode of inheritance and varying the minor allele frequency and the percentage of explained variance. Power comparison was then conducted with the most recently proposed MAX and allelic tests.

Finally, for associated loci, we aim to demonstrate that the GRM allows to formally test the underlying genetic model.

## 2 | MATERIALS AND METHODS

### 2.1 | General regression model

The GRM has been previously described for case-control design (Dizier et al., 2017). Briefly, the GRM allows testing for association without making assumptions on the genetic model, by using additive and dominance (deviation from additivity) terms, as proposed by Fisher and Wilson (Wilson, 1980).

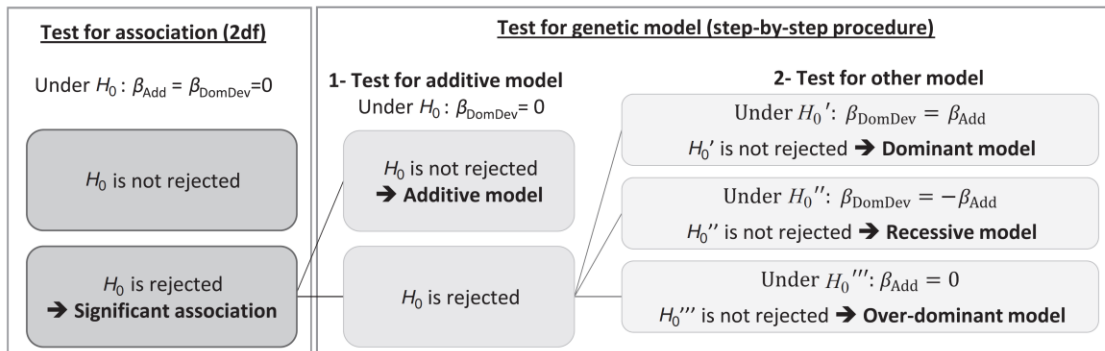
For quantitative traits, the LR model is written as:

$$P = \alpha + \beta_{\text{Add}} \cdot \text{Add} + \beta_{\text{DomDev}} \cdot \text{DomDev} + \beta_{\text{COV}} \cdot \text{COV} + E$$

$P$  is the quantitative phenotype,  $\beta_{\text{Add}}$  the regression coefficient for the additive effect,  $\beta_{\text{DomDev}}$  the regression coefficient for the dominance term, and  $E$  the error term. GRM is a particular “genotype effect model,” using the following coding scheme: 0/1/2 for the additive term and 0/1/0 for the dominance term, for AA, Aa, and aa genotypes, respectively. The association test is a likelihood ratio test. Under  $H_0$ ,  $\beta_{\text{Add}} = \beta_{\text{DomDev}} = 0$  and the statistic follows a chi-square distribution with two degrees of freedom.



$$P = \alpha + \beta_{\text{Add}} \cdot \text{Add} + \beta_{\text{DomDev}} \cdot \text{DomDev} + \beta_{\text{COV}} \cdot \text{COV} + E$$



**FIGURE 1** Statistical decisional diagram to test the genetic model using the general regression model (GRM)

In the case of significant association, the genetic models are tested by using a step-by-step procedure of one degree of freedom tests, as previously described (Dizier et al., 2017) (see Figure 1 for details).

First the additive model (under  $H_0: \beta_{\text{DomDev}} = 0$ ) is tested. If  $H_0$  is not rejected, the additive model is retained. If  $H_0$  is rejected, the dominant model (under  $H_0': \beta_{\text{DomDev}} = \beta_{\text{Add}}$ ) is first tested (i.e.,  $\beta_{\text{DomDev}}$  value is on the confidence interval defined by  $[\beta_{\text{Add}} \pm Z_{\alpha/2} \times \sigma_{\text{Add}}]$ , where  $z_{\alpha/2}$  is the critical value of the standard normal distribution with a probability of type I error  $\alpha$ ). If  $H_0'$  is not rejected, the dominant model is retained.

If  $H_0'$  is rejected, the recessive model (under  $H_0'': \beta_{\text{DomDev}} = -\beta_{\text{Add}}$ ) is tested. If  $H_0''$  is not rejected, the recessive model is retained. If  $H_0''$  is rejected, the overdominant model (under  $H_0''': \beta_{\text{Add}} = 0$ ) is tested. If  $H_0'''$  is not rejected, the overdominant (i.e., advantage of heterozygotes comparing to homozygotes) is retained. The regression models for each case are detailed in Table 1.

## 2.2 | Linear regression model

The LR is defined by the following formula:  $Y = \alpha + \beta \cdot X + \beta_{\text{COV}} \cdot \text{COV} + E$ . Different modes of transmission can be considered according to the coding scheme of  $X$  for AA, Aa, and aa genotypes, respectively:  $X$  is equal to 0, 1, and 2 in the case of an additive genetic model (LR\_ADD), 0, 0, 1 for a recessive model (LR\_REC) and 0, 1, 1 for a dominant model (LR\_DOM). The association test (under  $H_0, \beta = 0$ ) is a likelihood ratio test, following a one degree of freedom chi-square distribution.

## 2.3 | MAX test

The MAX test (Loley et al., 2013) is based on the generalized linear model. Specific genetic models can be modeled by introducing restrictions in terms of linear contrasts on the

regression parameter matrix. By using three dummy variables, the following contrasts were proposed for AA, Aa, and aa genotypes, respectively:  $-1, 0, 1$  for an additive genetic mode of inheritance,  $-0.5, -0.5, 1$  for a dominant mode, and  $-1, 0.5, 0.5$  for a recessive mode.

Association is tested using a classical Wald test on the regression parameter, under the three different genetic modes of inheritance (additive, dominant, and recessive). The MAX test allows selecting the most plausible genetic model using the minimum  $P$ -value of the three association tests.

## 2.4 | Allelic test

The allelic test proposed for quantitative traits (Lee et al., 2013) is an extension of the allele-based test used for case-control studies, using the alleles as statistical units instead of genotypes, as done in the LR approach.

## 2.5 | Simulations

For power comparison between GRM and LR, a total of 100,000 replicates of samples of 1,000 individuals were simulated. A quantitative trait was generated, following a standard normal distribution and assuming a single causal variant. Various genetic models were considered, by using four minor allele frequencies (MAFs of 0.1, 0.2, 0.3, and 0.4), several percentages of explained variance (PEVs, varying between 1% and 15%), and the three classical genetic models (additive, dominant, and recessive). To assess the sample size effect, another set of 100,000 replicates of 5,000 individuals was simulated, for the same MAFs and PEVs.

Because of the intensive computational time of the MAX test, which was 22-fold more computationally expensive than GRM, a subset of 1,000 replicates was used to compare MAX and GRM tests, for all MAFs and PEVs simulated for 1,000 individuals. The same subset was used for power comparison with the allelic test. To estimate the type I error rate of GRM



**TABLE 1** Quantitative regression models according to the genetic models and the three genotypes

Models	Genotypes		
	AA	Aa	aa
General	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + \beta_{\text{Add}} + \beta_{\text{DomDev}} + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + 2\beta_{\text{Add}} + \beta_{\text{COV}} \cdot \text{COV} + E$
Additive	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + \beta_{\text{Add}} + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + 2\beta_{\text{Add}} + \beta_{\text{COV}} \cdot \text{COV} + E$
Dominant	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + 2\beta_{\text{Add}} + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + 2\beta_{\text{Add}} + \beta_{\text{COV}} \cdot \text{COV} + E$
Recessive	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + 2\beta_{\text{Add}} + \beta_{\text{COV}} \cdot \text{COV} + E$
Overdominant	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + \beta_{\text{DomDev}} + \beta_{\text{COV}} \cdot \text{COV} + E$	$P = \alpha + \beta_{\text{COV}} \cdot \text{COV} + E$

and LR tests, 1.0E8 replicates were simulated under the null hypothesis of no association (under  $H_0$ , PEV is null).

## 2.6 | Association tests

Association analyses were performed for all simulated replicates using LR\_ADD, LR\_DOM, LR\_REC, and GRM tests. A subset of simulated replicates was analyzed to compare GRM, MAX, and allelic tests. The normality of the residual was checked for each regression model. The threshold of 5.0E-8, consensually used for GWAS of quantitative traits (Risch & Merikangas, 1996), was applied to declare significant association for all tests (LR\_, GRM, MAX, and allelic tests).

The type I error rate was estimated by the percentage of replicates showing significant association. The empirical power for GRM, LR, MAX, and allelic tests was estimated as the percentage of simulated replicates showing significant association.

On real data, the significance level for association analysis was fixed by applying the Bonferroni correction on the number of independent tests. This number corresponds to the total number of effective single-nucleotide polymorphism (SNP) tests, after filtering out the correlation between SNP tests caused by linkage disequilibrium (LD) between SNPs (Gao, Starmer, & Martin, 2008).

## 2.7 | Test of the genetic model by the GRM

The test of genetic models is only applied to SNPs found significantly associated to the disease. The significance threshold considered to test the genetic model was corrected for multiple testing by applying Bonferroni correction (i.e., equal to 5% divided by the number of SNPs considered for model testing). For simulation study, we arbitrarily applied a threshold at 0.01 but another threshold could be considered.

## 2.8 | Genetic model retained by the MAX test

Although the MAX test does not formally test the underlying genetic model, it can be defined by the model having the smallest  $P$ -value for MAX association tests (Loley et al., 2013).

## 2.9 | Commands and shell scripts

Simulations, association tests, and genetic model testing were conducted using the PLINK (Purcell et al., 2007) and R software (R Core Team, 2017). PLINK commands and shell scripts are available in the Supporting Information section of the additional files.

## 2.10 | Cohort description

The D.E.S.I.R. (data from an epidemiological study on the insulin resistance syndrome) cohort has been described elsewhere (Balkau, 1996). Briefly, participants were recruited from volunteers insured by the French social security system, which offered periodic health examinations free of charge. Volunteers were recruited as part of the 9-year follow-up study, to study the insulin resistance syndrome. All participants signed an informed consent and the protocol was approved by the Bicêtre Hospital ethics committee. The selection criteria for the present study were to be at least 45 years old, have a body mass index (BMI) less than 27 kg/m<sup>2</sup>, and have an HbA1c level lower than 6.5%, to exclude people with diabetes (ADA, 2017). Participants were genotyped by the Illumina HumanHap300 beadchip.

Genotyping was duplicated on two different platforms and genotypes were obtained for an average of 99.4% of markers for each subject with a reproducibility of >99.9%. Genotyping and quality control have been previously described (Bouatia-Naji et al., 2008). Exclusion criteria for SNPs were (1) less than 5% of missing data, (2) significant Hardy–Weinberg equilibrium test ( $P$ -value lower than 1.0E-4), and (3) a MAF lower than 5%.

A principal component analysis was performed to adjust for population stratification. The first 10 principal components were included in the analysis as recommended (Reed et al., 2015).

## 3 | RESULTS

### 3.1 | Type I error rate

Under the null hypothesis of no association, the estimated type I error rate was close to the theoretical thresholds of 5%, 1%,

and  $1.0E-5$ . GRM robustness was similar to those of the three LR models (see additional Table S1 in Supporting Information for details).

### 3.2 | Comparison of power to detect association between GRM and LR tests

Results for samples of 1,000 individuals are provided in Figure 2. Results were not shown for a PEV greater than 10%, as the power difference was null because all powers reached 100%.

For the additive simulated model, the power of GRM to detect association was close to LR\_ADD one, for all MAF and PEV tested. For a PEV lower or equal to 5%, LR\_ADD was more powerful than GRM, but the power difference never exceeded 8.6% (Figure 2 and Supporting Information Figure S1).

For the dominant simulated model, GRM was as or more powerful than LR\_ADD to detect association. For a PEV lower or equal to 5%, the power difference increased with MAF (Figure 2 and Supporting Information Figure S1) and the highest gain of power using GRM is observed for an MAF of 0.4 (24.1%). As expected, the LR\_DOM test always had the highest power when the simulated model was dominant, but the difference of power compared to GRM never exceeded 8.7%.

For the recessive simulated model, GRM was always as or more powerful than the LR\_ADD test. This gain of power increased when the MAF decreased, reaching 80% for a PEV of 5% and an MAF of 0.1 or 0.2 (Figure 2 and Supporting Information Figure S1). As expected, the LR\_REC test was as or more powerful than the GRM when the simulated model was recessive, but the difference in the GRM never exceeded 3.4%.

Results for 5,000 individuals are provided in Supporting Information Figures S2 and S3. Results are similar to those for 1,000 individuals, with a gain of power using GRM when the simulated model was dominant and a greater gain when it was recessive. Note that the maximum gain of power using a larger sample size (5,000 instead of 1,000) was observed for smaller PEV (equal to 1% instead of 5%), with a gain of power of 16% under a dominant model and of 88% under a recessive model.

### 3.3 | Tests of the genetic model by GRM

Results for samples of 1,000 individuals are provided in Figure 3. For an additive simulated model, the true model was retained in most replicates. For an MAF greater than 0.1, the proportion of replicates retaining the true model was close to 1 (type I error) and was always greater than 96%. For an MAF equal to 0.1, this proportion ranged between 66% and 98%. For a dominant simulated, the percentage of replicates

concluding to the true model increased with increasing PEV and the MAF, ranging between 70% and 82% for an MAF greater than 0.1 and a PEV greater than 5%. For an MAF equal to 0.1, this percentage decreased between 32% and 55% for a PEV greater than 5% and became very low (less than 10%) for a PEV smaller than 5%, because of the lack of power of the association test. For a recessive simulated model, the percentage of replicates concluding to the true model ranged between 89% and 100%. This percentage decreased when the MAF increased (100% for an MAF of 0.1 versus ranging between 89% and 92% for an MAF of 0.4).

For a sample of 5,000 individuals, similar results were obtained. The percentage of replicates retaining the true simulated model was higher with 5,000 than with 1,000 individuals (see Supporting Information Figure S4 in supplementary data for details).

### 3.4 | Power comparison of GRM and MAX test to detect association

Results of power comparison are detailed in Figure 4 and Supporting Information Table S2. The GRM test was always as or more powerful than the MAX test and particularly more powerful for a dominant and additive genetic mode of inheritance for an MAF of 0.1 (for a PEV of 5%, power difference reaching 67% and 91% for an additive and dominant mode, respectively).

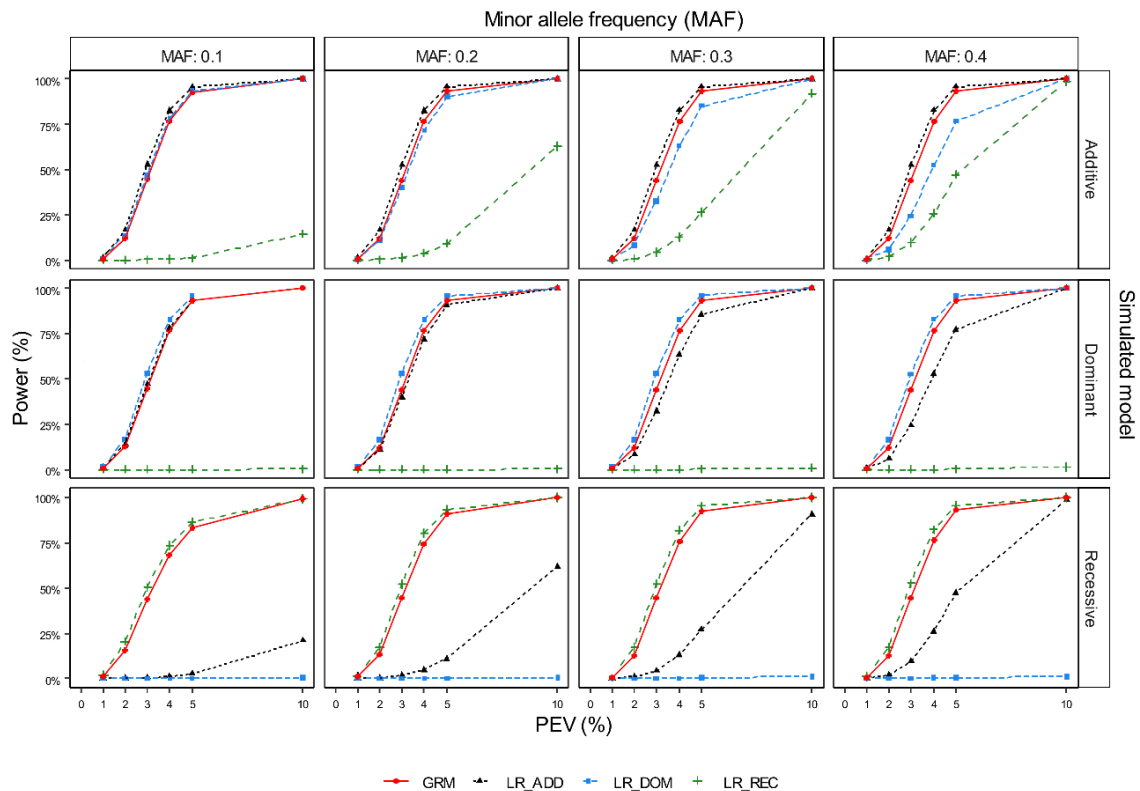
### 3.5 | Comparison of GRM and MAX tests to define the underlying genetic mode of transmission

Figure 4 summarizes the proportion of replicates retaining an additive, recessive, or dominant genetic model by the GRM and the MAX tests according to the simulated genetic model and by varying the mode of transmission, the minor allele frequencies, and the percentage of explained variance.

For an additive simulated model, the GRM test retained the true model for most replicates, whereas the MAX test selected the dominant genetic model as the most plausible. For a dominant simulated model, both GRM and MAX tests retained the dominant model in most replicates. For a recessive simulated model and low MAFs ( $\leq 0.2$ ), GRM has always a greater sensitivity to retain the true genetic model. For higher MAFs ( $\geq 0.3$ ), the difference between GRM and MAX was weaker meaning that both tests have a similar ability to select the true mode of inheritance.

### 3.6 | Power comparison of GRM and allelic test to detect association

Results are detailed in Supporting Information Table S2. For an additive simulated model, the GRM and the allelic tests



**FIGURE 2** Power to detect association of the general regression model (GRM) and linear regression (LR) tests, for a  $P$ -value threshold of  $5E-8$  for the association test, by varying mode of transmission, the minor allele frequencies, and the percentage of explained variance and for a sample of 1,000 individuals [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

have similar power. For a dominant simulated model, they also have similar power, except for a PEV of 5% and an MAF greater than 0.2, where the GRM was more powerful than the allelic test. For a recessive model, the GRM was even more powerful than the allelic test (power difference reaching to 79.5% for an MAF of 0.1 and a PEV of 5%).

### 3.7 | Application to the D.E.S.I.R. cohort

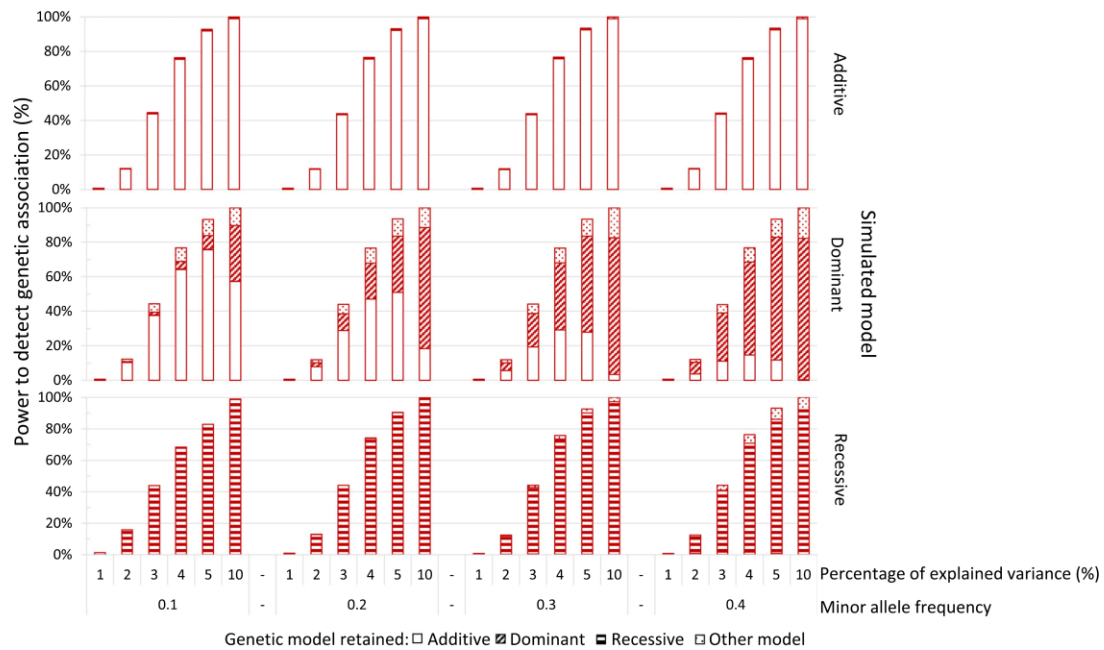
As part of the D.E.S.I.R. cohort, a sample of 695 healthy volunteers having an HbA1c level lower than 6.5% was included in the present study (41% men and 59% women), with a mean age of 53.5 ( $\pm 5.6$ ) years and a mean BMI of 23.2 ( $\pm 1.8$ ) kg/m<sup>2</sup>. The mean HbA1c level at inclusion was 5.26% ( $\pm 0.36\%$ ). No gender effect on HbA1c was observed. To account for center and age effects and as the age effect was different according to the center, HbA1c levels were adjusted before association testing.

After quality control, 301,183 SNPs were analyzed. PCA analysis showed that only one population was present in the D.E.S.I.R. cohort (see Supporting Information Figure S5). Script shells to apply GRM to real data are provided in

the Supporting Information section of the additional files. Results are provided in Table 2 and in Supporting Information Figure S6. Using the Bonferroni correction applied to the number of independent tests ( $N = 147,360$ ) estimated using the Gao method (Gao et al., 2008), the  $P$ -value threshold for significance was fixed at  $3.4E-7$ . Two SNPs were associated to HbA1c: rs6468861, located on 8q22.3, significantly associated using LR\_ADD ( $P$ -value =  $2.5E-7$ ), and suggestively associated using GRM ( $P$ -value =  $1.3E-6$ ) and rs4594032, located on the 11q21 chromosomal region, significantly associated using GRM ( $P$ -value =  $1.8E-7$ ), but not using LR\_ADD ( $P$ -value = 0.02). As shown in the Supporting Information Figure S7, results remain unchanged after taking into account population stratification. No other SNP has been detected using LR\_REC and LR\_DOM tests.

Results of genetic model testing are shown in Table 3. For rs6468861,  $\beta_{\text{DomDev}}$  was not significantly different from zero ( $P$ -value = 0.87), so the additive model of inheritance was retained. A dose-effect of the T allele was observed on the HbA1c level. Compared to CC genotypes, mean HbA1c level was increased by 0.09 for TC genotypes and by 0.18 for TT





**FIGURE 3** Proportion of replicates significantly associated ( $P$ -value threshold of  $5E-8$ ) retaining the additive, dominant, or recessive model by using the general regression model (GRM) tests and using a threshold of 0.01, according to the simulated model and varying the mode of transmission, the minor allele frequency and the percentage of explained variance and for a sample of 1,000 individuals [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 2** HbA1c genome-wide association studies (GWAS) results on the D.E.S.I.R. sample

SNP	Position	Allele	MAF		P-value			
			CEU <sup>a</sup>	D.E.S.I.R.	LR_ADD	LR_DOM	LR_REC	GRM
rs6468861	8q22.3	T/C	0.37	0.42	2.5E-7	6.9E-6	1.5E-4	<i>1.3E-6</i>
rs4594032	11q21	G/T	0.44	0.41	0.016	4.5E-6	0.17	<b>1.8E-7</b>

D.E.S.I.R., Data from an epidemiological study on the insulin resistance syndrome; GRM, general regression model; LR\_ADD, linear regression under additive genetic model; LR\_DOM, linear regression under dominant genetic model; LR\_REC, linear regression under recessive genetic model; MAF, minor allele frequency; SNP, single-nucleotide polymorphism; bold, significant  $P$ -value ( $P < 3.4E-7$ ); italic, suggestive  $P$ -value ( $P < 3.4E-6$ ).

<sup>a</sup>Minor allele frequencies are based on Thousand Genomes CEU (Northern Europeans from Utah, May 2011).

**TABLE 3** Test of genetic model on the D.E.S.I.R sample

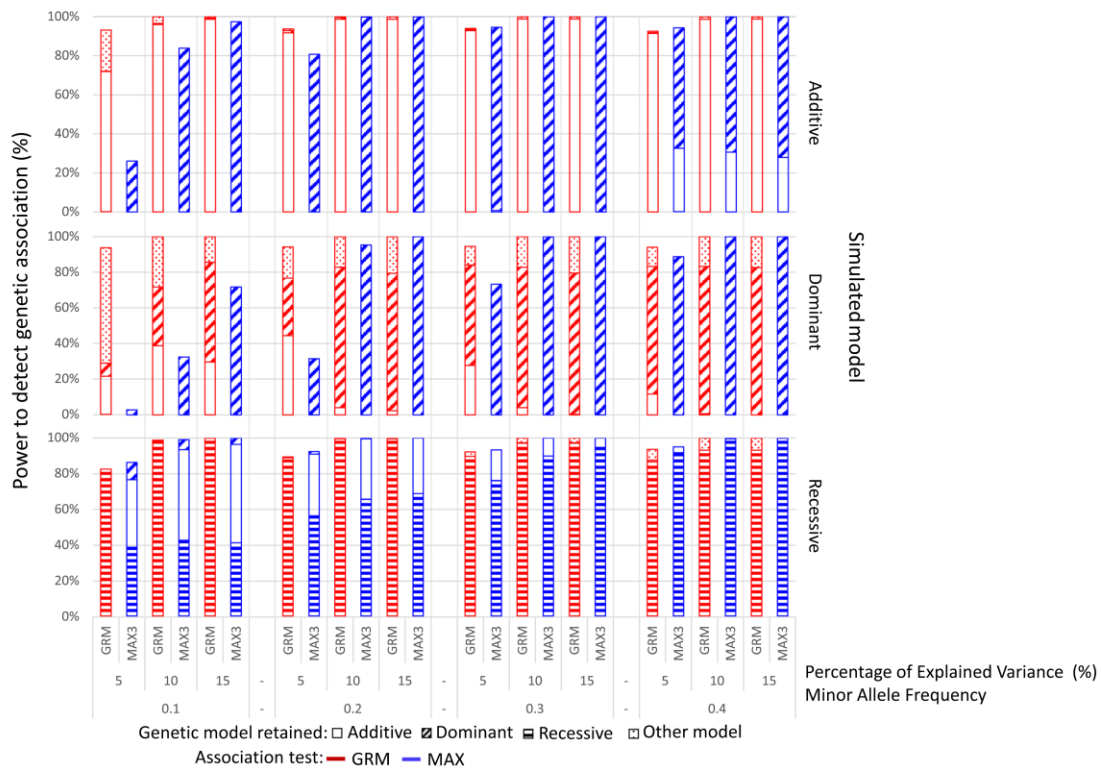
SNP	$\beta_{Add}$	$\beta_{Add}$ P-value	IC <sub>97.5%</sub> ( $\beta_{Add}$ )	$\beta_{DomDev}$	$\beta_{DomDev}$ P-value	Genetic model
rs6468861	0.09	7.1E-7	(0.05;0.13)	0.004	0.87	Additive
rs4594032	-0.02	0.28	(-0.06;0.02)	-0.13	6.9E-7	Overdominant

D.E.S.I.R., Data from an epidemiological study on the insulin resistance syndrome.

genotypes, respectively. For rs4594032, the additive model was rejected (i.e.,  $\beta_{DomDev}$  was significantly different from zero;  $P$ -value =  $6.9E-7$ ) and  $\beta_{Add}$  was not significantly different from zero ( $P$ -value = 0.28), so the overdominance model was retained. Mean HbA1c level was decreased by 0.13 in heterozygotes (GT genotype) compared to homozygotes (GG and TT genotypes). Percentage of explained variance under the genetic model retained were, respectively, 3.6% for rs6468861 and 4% for rs4594032.

## 4 | DISCUSSION

As for binary traits (Dizier et al., 2017), we propose to use the GRM test for GWAS of quantitative traits, as its power to detect associations did not depend on the underlying mode of transmission. An important finding is that the difference of power between GRM and LR\_ADD tests is not very important when the true model is additive although the GRM test has one supplementary degree of freedom compared to the



**FIGURE 4** Power comparison between general regression model (GRM) and MAX tests using a  $P$ -value threshold of  $5E-8$  and proportion of replicates retaining additive, recessive, or dominant genetic model according to the simulated genetic model and by varying the mode of transmission, the minor allele frequencies, and the percentage of explained variance [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

LR\_ADD test. Indeed, in case of equivalent models, addition of a supplementary degree of freedom would conduct to a less powerful test.

Moreover, the GRM may be more powerful than the LR\_ADD test when the true model was dominant and much more when the true model was recessive (gain of power reaching 80%). Additionally, GRM and LR\_ tests had similar computation time. Our results are similar of those previously published by Lettre et al. for a PEV of 1% (Lettre et al., 2007). Considering greater PEVs, gain of power of the GRM comparing to LR\_ADD were observed, not only for a recessive simulated model but also for a dominant one. In a same manner, when comparing the GRM to the allelic test, similar power was observed when the underlying genetic model is additive but not when there is departure from additivity. In this case, the GRM association test is more powerful than the allelic test. Concerning the MAX test, our results confirmed those of Loley et al. (2013) showing a good power of the MAX test for minor allele frequencies equal or greater than 0.3 and, in this case, similar to the power of GRM but, for lower minor allele frequencies, the GRM was more powerful than the MAX test.

The underlying model is not formally tested by the MAX test. The retained genetic model corresponds to the genetic model (here, additive, dominant, or recessive) having the lowest adjusted  $P$ -value (Kitsche et al., 2016; Konietzschke, Libiger, & Hothorn, 2012; Loley et al., 2013). The GRM had a similar or a greater ability to retain the true simulated model than the MAX test, especially under an additive model.

GRM can be considered as another formulation of the two degrees of freedom genotypic test, but our specific coding scheme is the only one allowing to formally test the underlying genetic model (dominant, recessive, additive, overdominant, etc.). However, the GRM has some limitations. For power comparison between GRM and LR\_ or MAX tests, our simulations were done considering the simulated SNP as the disease locus (i.e., with a complete LD between the SNP and the disease locus of the quantitative trait). Several tests have been developed to take into account LD (Fan & Xiong, 2002; Fan, Jung, & Jin, 2006; Nielsen & Weir, 1999, 2001) but they have not been compared to GRM.

The GRM association test is based on the regression method. In case of departure from the assumptions for regression test (i.e., normality, linearity, homoscedasticity, and

absence of multicollinearity of the residuals), a nonparametric association test would be preferable. However, the most currently used nonparametric tests are not model free and so sensitive to departure from the genetic model (Jonckheere, 1954; Terpstra, 1952) and are also time consuming.

Application of GRM to the D.E.S.I.R. data well illustrated the gain in power of the GRM to detect associations compared to the LR\_ADD when there is deviation from additivity: rs4594032 was significantly associated by using the GRM test but not by using the LR\_ADD test. This SNP showed deviation from additivity. It had been slightly suggested as associated ( $P$ -value = 0.007) in the Diabetes Genetics Replication and Meta-Analysis (DIAGRAM) Consortium (Mahajan et al., 2014). The LD region of rs4594032 (using an  $r^2$  threshold of 0.6) is chr11:96467132-96535365 (hg19), no SNP in LD with rs459032 was genotyped in this region in our sample. The closest gene in the LD region is JRKL (for jerky homolog-like), located at 343 kb of rs4594032, which is homologous in rat, on chromosome 8, to a region linked to insulin resistance (Marissal-Arvy, Helies, Tridon, Moisan, & Mormede, 2014). Two regions of open chromatin and one region of enhancer fixation are identified in the LD region. Interestingly, the SNP rs4594032 is located in the FGQTL3 region for fasting plasma glucose level QTL3 (11q21-q22; 11: 93,000,000–110,600,000), including the *MNTR1B* gene known to be associated with type 2 diabetes (Bonnefond et al., 2012; Bouatia-Naji et al., 2009; Dupuis et al., 2010; Lyssenko et al., 2009), fasting glucose (Bouatia-Naji et al., 2009; Dupuis et al., 2010; Prokopenko et al., 2009), and the residual function of the  $\beta$  cells (Dupuis et al., 2010; Lyssenko et al., 2009).

The other SNP rs6468861 found to be associated with HbA1c by both LR\_ADD and GRM is located in the intron of the gene *BAALC* (for brain and acute leukemia cytoplasmic), known to be associated with leukemia and suggested to have a synaptic role in postsynaptic lipid rafts. The LD region (defined by  $r^2 > 0.6$ ), is chr8:104129424-104181000 (Hg19). No SNP in LD with rs6468861 was genotyped in this region in our sample. The LD region contains a quantitative trait locus controlling the expression level of fetal hemoglobin (chr8: 45200001-145138636) (Garner, Tatu, Best, Creary, & Thein, 2002).

Overall, the GRM test is a user-friendly and powerful genetic test to detect association and determine the underlying model as well as for quantitative than for binary traits. This test is easy and quick to apply and thus particularly appropriate for association studies of large panels of markers.

## LIST OF ABBREVIATIONS

BMI: body mass index; COV: covariates; D.E.S.I.R.: data from an epidemiological study on the insulin resistance

syndrome; DIAGRAM: diabetes genetics replication and meta-analysis consortium; E: error term; GWAS: genome-wide association studies; GRM: general regression model; HbA1c: glycated hemoglobin; LR: linear regression model; LR\_ADD (additive), LR\_DOM (dominant), LR\_REC (recessive); MAFs: minor allele frequencies; MAX test: maximum of the standardized version of optimal tests; MAX: max test developed by Loley et al., 2013; NMAX: nonparametric versions of the MAX test; PEV: proportion of explained variance; QTL: quantitative trait loci.

## ACKNOWLEDGMENTS

We are grateful to all participants of this D.E.S.I.R. cohort study (Inserm U1018: Beverley Balkau, Pierre Ducimetière, Eveline Eschwège; Inserm U1138: François Alhenc-Gelas, Frédéric Fumeron, Michel Marre, Ronan Roussel; CHU d'Angers: Yves Gallois, Andrée Girault; CHU de Rennes: Fabrice Bonnet; CNRS UMR8090: Amélie Bonnefond, Stéphane Cauchi, Philippe Froguel; Centres d'examen de santé de l'Assurance Maladie: Alençon, Angers, Blois, Caen, Chartres, Chateauroux, Cholet, Le Mans, Orléans, Tours; Institut de Recherche en Médecine Générale: Joël Cogneau; General practitioners of the region; Institut inter-Régional pour la Santé (IRSA): Catherine Born, Emile Caces, Martine Cailleau, Nane Copin, Olivier Lantieri, Jean-Guy Moreau, Fabienne Rakotozafy, Jean Tichet, Sylviane Vol). We are grateful to Dr. Andreas Ziegler and Dr. Christina Loley for providing the R-script of the MAX test dedicated to GWAS. We are also grateful to the INRA MIGALE bioinformatics platform (<http://migale.jouy.inra.fr>) for providing computational resources.

## CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

## AUTHOR CONTRIBUTIONS

F.M. and M.-H.D. jointly conceived the study and designed the simulation models. E.G. implemented the simulation study and conducted all simulation data analyses. E.G., F.M., and M.-H.D. interpreted the results. E.G., F.M., M.-H.D., and B.B. wrote the paper. M.B. conducted analysis on D.E.S.I.R. data. P.F., G.R., and M.C. provided access to D.E.S.I.R. genetic data. J.T., R.R., and B.B. provided access to D.E.S.I.R. phenotypic data. C.J. heads the UMR-S958 Inserm unit. All authors discussed the results and implications and commented on the manuscript at all stages. All authors gave approval of the version to be published.



## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

We obtained the D.E.S.I.R. data by contacting the members of the D.E.S.I.R. group and all participants signed an informed consent. The protocol was approved by the Bicêtre Hospital ethics committee.

## DATA AVAILABILITY STATEMENT

The datasets generated during the current study are not publicly available because of the very large size of the simulated data (216 models of 100,000 replicates) but script to replicate simulation analysis and D.E.S.I.R. study are available in the Supporting Information.

## ORCID

Emilie Gloaguen  <https://orcid.org/0000-0002-5205-2262>

Mathilde Boissel  <https://orcid.org/0000-0001-5719-8749>

Flavie Mathieu  <https://orcid.org/0000-0002-3386-0299>

## REFERENCES

- ADA. (2017). Classification and diagnosis of diabetes. *Diabetes Care*, 40(Suppl 1), S11–S24. Retrieved from <https://doi.org/10.2337/dc17-S005>
- Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, 7(10), 781–791. Retrieved from <https://doi.org/10.1038/nrg1916>
- Balkau, B. (1996). An epidemiologic survey from a network of French health examination centres (D.E.S.I.R.): Epidemiologic data on the insulin resistance syndrome. *Revue d'Epidémiologie et de Santé Publique*, 44(4), 373–375.
- Bonnefond, A., Clement, N., Fawcett, K., Yengo, L., Vaillant, E., Guillaume, J. L., ... Froguel, P. (2012). Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nature Genetics*, 44(3), 297–301. Retrieved from <https://doi.org/10.1038/ng.1053>
- Bouatia-Naji, N., Rocheleau, G., Van Lommel, L., Lemaire, K., Schuit, F., Cavalcanti-Proenca, C., ... Froguel, P. (2008). A polymorphism within the *G6PC2* gene is associated with fasting plasma glucose levels. *Science*, 320(5879), 1085–1088. Retrieved from <https://doi.org/10.1126/science.1156849>
- Bouatia-Naji, N., Bonnefond, A., Cavalcanti-Proenca, C., Sparso, T., Holmkvist, J., Marchand, M., ... Froguel, P. (2009). A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nature Genetics*, 41(1), 89–94. Retrieved from <https://doi.org/10.1038/ng.277>
- Dizier, M. H., Demenais, F., & Mathieu, F. (2017). Gain of power of the general regression model compared to Cochran–Armitage trend tests: Simulation study and application to bipolar disorder. *BMC Genetics*, 18(1), 24. Retrieved from <https://doi.org/10.1186/s12863-017-0486-6>
- Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A. U., ... Barroso, I. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics*, 42(2), 105–116. Retrieved from <https://doi.org/10.1038/ng.520>
- Fan, R., & Xiong, M. (2002). High resolution mapping of quantitative trait loci by linkage disequilibrium analysis. *European Journal of Human Genetics*, 10(10), 607–615. Retrieved from <https://doi.org/10.1038/sj.ejhg.5200843>
- Fan, R., Jung, J., & Jin, L. (2006). High-resolution association mapping of quantitative trait loci: A population-based approach. *Genetics*, 172(1), 663–686. Retrieved from <https://doi.org/10.1534/genetics.105.046417>
- Freidlin, B., Zheng, G., Li, Z., & Gastwirth, J. L. (2002). Trend tests for case-control studies of genetic markers: Power, sample size and robustness. *Human Heredity*, 53(3), 146–152. Retrieved from <https://doi.org/64976>
- Gao, X., Starmer, J., & Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology*, 32(4), 361–369. Retrieved from <https://doi.org/10.1002/gepi.20310>
- Garner, C. P., Tatu, T., Best, S., Creary, L., & Thein, S. L. (2002). Evidence of genetic interaction between the  $\beta$ -globin complex and chromosome 8q in the expression of fetal hemoglobin. *American Journal of Human Genetics*, 70(3), 793–799. Retrieved from <https://doi.org/10.1086/339248>
- Jonckheere, A. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, 41, 133–145.
- Kitsche, A., Ritz, C., & Hothorn, L. A. (2016). A general framework for the evaluation of genetic association studies using multiple marginal models. *Human Heredity*, 81(3), 150–172. Retrieved from <https://doi.org/10.1159/000448477>
- Konietschke, F., Libiger, O., & Hothorn, L. A. (2012). Nonparametric evaluation of quantitative traits in population-based association studies when the genetic model is unknown. *PLoS ONE*, 7(2), e31242. Retrieved from <https://doi.org/10.1371/journal.pone.0031242>
- Lee, S. M., Karrison, T. G., Cox, N. J., & Im, H. K. (2013). Quantitative allelic test—A fast test for very large association studies. *Genetic Epidemiology*, 37(8), 831–839. Retrieved from <https://doi.org/10.1002/gepi.21768>
- Lette, G., Lange, C., & Hirschhorn, J. N. (2007). Genetic model testing and statistical power in population-based association studies of quantitative traits. *Genetic Epidemiology*, 31(4), 358–362. Retrieved from <https://doi.org/10.1002/gepi.20217>
- Loley, C., Konig, I. R., Hothorn, L., & Ziegler, A. (2013). A unifying framework for robust association testing, estimation, and genetic model selection using the generalized linear model. *European Journal of Human Genetics*, 21(12), 1442–1448. Retrieved from <https://doi.org/10.1038/ejhg.2013.62>
- Lyssenko, V., Nagorny, C. L., Erdos, M. R., Wierup, N., Jonsson, A., Spiegel, P., ... Groop, L. (2009). Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nature Genetics*, 41(1), 82–88. Retrieved from <https://doi.org/10.1038/ng.288>
- Mahajan, A., Go, M. J., Zhang, W., Below, J. E., Gaulton, K. J., Ferreira, T., ... Morris, A. P. (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature Genetics*, 46(3), 234–244. Retrieved from <https://doi.org/10.1038/ng.2897>
- Marissal-Arvy, N., Helies, J. M., Tridon, C., Moisan, M. P., & Mormede, P. (2014). Quantitative trait loci influencing abdominal fat deposition and functional variability of the HPA axis in the rat. *Hormone and Metabolic Research*, 46(9), 635–643. Retrieved from <https://doi.org/10.1055/s-0034-1383574>

- McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P., & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9(5), 356–369. Retrieved from <https://doi.org/10.1038/nrg2344>
- Nielsen, D. M., & Weir, B. S. (1999). A classical setting for associations between markers and loci affecting quantitative traits. *Genetic Research*, 74(3), 271–277. Retrieved from <https://doi.org/10.1017/s0016672399004231>
- Nielsen, D. M., & Weir, B. S. (2001). Association studies under general disease models. *Theoretical Population Biology*, 60(3), 253–263. Retrieved from <https://doi.org/10.1006/tpbi.2001.1539>
- Prokopenko, I., Langenberg, C., Florez, J. C., Saxena, R., Soranzo, N., Thorleifsson, G., ... Abecasis, G. R. (2009). Variants in MTNR1B influence fasting glucose levels. *Nature Genetics*, 41(1), 77–81. Retrieved from <https://doi.org/10.1038/ng.290>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. Retrieved from <https://doi.org/10.1086/519795>
- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org>
- Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., & Foulkes, A. S. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Statistics in Medicine*, 34(28), 3769–3792. Retrieved from <https://doi.org/10.1002/sim.6605>
- Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, 273(5281), 1516–1517.
- Schaid, D. J., McDonnell, S. K., Hebring, S. J., Cunningham, J. M., & Thibodeau, S. N. (2005). Nonparametric tests of association of multiple genes with human disease. *American Journal of Human Genetics*, 76(5), 780–793. Retrieved from <https://doi.org/10.1086/429838>
- Slager, S. L., & Schaid, D. J. (2001). Case-control studies of genetic markers: Power and sample size approximations for Armitage's test for trend. *Human Heredity*, 52(3), 149–153. Retrieved from <https://doi.org/53370>
- So, H. C., & Sham, P. C. (2011). Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. *Behavior Genetics*, 41(5), 768–775. Retrieved from <https://doi.org/10.1007/s10519-011-9450-9>
- Terpstra, T. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae*, 14, 327–333.
- Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*, 90(1), 7–24. Retrieved from <https://doi.org/10.1016/j.ajhg.2011.11.029>
- Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., ... Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, 42(Database issue), D1001–D1006. Retrieved from <https://doi.org/10.1093/nar/gkt1229>
- Wilson, S. R. (1980). A note on the correct definition of additive deviation and dominance deviation. *Annals of Human Genetics*, 44(1), 113–115. Retrieved from <https://doi.org/10.1111/j.1469-1809.1980.tb00951.x>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Gloaguen E, Dizier M-H, Boissel M, et al. General regression model: A “model-free” association test for quantitative traits allowing to test for the underlying genetic model. *Ann Hum Genet.* 2019;1–11. <https://doi.org/10.1111/ahg.12372>



# Chapitre 3 | Facteurs influençant le taux d'HbA1c chez le patient souffrant de DT1 – Revue de la littérature

---

## 3.1 Résumé

Il existe un grand nombre de facteurs qui influencent le taux d'HbA1c chez les sujets non-diabétiques, aussi bien cliniques que génétiques. L'étude de ces facteurs chez les patients souffrant de DT1 est plus complexe, puisque le traitement à l'insuline influence directement le taux d'HbA1c. De nombreuses études ont porté sur le taux d'HbA1c chez les enfants souffrant de DT1 lors de la première année suivant le diagnostic de diabète, période dite de rémission partielle et ont fait l'objet d'une revue (Mazarello Paes *et al.*, 2017). Cette revue portant sur 20 études a suggéré différents facteurs associés à une augmentation du taux d'HbA1c : avoir un âge de diagnostic supérieur à 10 ans, ne pas être caucasien, être dans une famille monoparentale, avoir un taux élevé d'HbA1c au diagnostic, ainsi que de bénéficier d'un traitement non-intensif (peu d'injections et pas de pompe). Ces facteurs doivent encore être confirmés par d'autres études.

D'autres études ont été conduites pour analyser l'impact de facteurs après deux ans de diabète, la plupart regardant uniquement l'effet d'un ou de quelques facteurs. Une revue des articles publiés entre janvier 1993 et décembre 2010 portant sur les facteurs démographiques et liés à l'individu a conclu à l'effet de certains facteurs présentés dans le Tableau 11 (Neylon, O'Connell, Skinner, & Cameron, 2013).

Depuis cette revue, de nombreuses études ont été publiées pour identifier de nouveaux facteurs, ce qui a justifié une nouvelle revue systématique de la littérature. Notre revue a porté sur les articles publiés entre janvier 2011 et juillet 2016 portant sur le taux d'HbA1c chez l'enfant, l'adolescent et le jeune adulte (âgés de moins de 25 ans) atteint de DT1 depuis plus d'un an.

Type de facteurs étudiés	Facteurs associés à une détérioration du taux d'HbA1c
Facteurs démographiques	Être adolescent Appartenir à un groupe ethnique minoritaire Vivre dans une structure familiale monoparentale Avoir un niveau socio-économique faible
Facteurs interpersonnels	Avoir peu de support familial Avoir beaucoup de conflits familiaux Avoir peu de surveillance et d'aide familial Avoir des mauvaises influences dans son cercle amical
Facteurs intra-personnels	Être peu consciencieux Avoir des désordres alimentaires ou des troubles psychiatriques
Facteurs technologiques	Faire peu d'autotests de glycémie par jour

**Tableau 11** : Facteurs identifiés comme étant associés à un moins bon équilibre glycémique (Neylon *et al.*, 2013)

Nous avons sélectionné les articles à partir de Ovid Medline (Pubmed et Web of Science) et de Cochrane Library à l'aide de mots-clefs et de critères de sélection précis. Chaque article a ensuite été classé selon un critère de qualité (mauvais :  $\leq 2$  ; intermédiaire : 3-4 ; et élevé :  $\geq 5$ ) en fonction de différents paramètres (puissance de l'étude, statistique utilisée, correction pour les tests multiples, ajustement pour les autres facteurs dont le traitement). Les articles ayant une qualité jugée mauvaise n'ont pas été inclus dans la revue.

Pour les facteurs non génétiques, nous avons sélectionné 78 articles sur les 1974 articles identifiés, dont 38 avec une qualité élevée et 40 avec une qualité intermédiaire. Les études identifiées sont très variables, avec des études de cohortes, des essais cliniques et des études transversales et portent sur des tailles d'échantillons très différentes (de 90 à plus de 30 000 patients souffrant de DT1), présentées en Annexe 2. De plus, le taux d'HbA1c a été analysé de façon différente d'une étude à l'autre : certaines études ont analysé le taux d'HbA1c de façon quantitative, d'autres en créant des catégories pour lesquelles les seuils permettant de catégoriser le taux d'HbA1c étaient différents. Dans leurs analyses, un certain nombre de ces études n'ont pas pris en compte de covariable, ni l'effet du traitement. En raison de ces différents points, nous n'avons donc pas pu effectuer de méta-analyse. Nous avons réalisé une analyse qualitative des principaux facteurs influençant le taux d'HbA1c ainsi que les interactions qui existaient entre eux.

Les principaux facteurs non-génétiques identifiés associés au taux d'HbA1c sont démographiques (l'âge, l'ethnie, le statut socio-économique), liés au DT1 (la durée du DT1, le type de traitement, la dose d'insuline), à l'adhérence ou à la personnalité du patient, avec notamment un impact de la présence de symptômes dépressifs ou anxieux. Un certain nombre de ces facteurs sont corrélés entre eux, ce qui rend l'analyse complexe. Par exemple, les patients ayant un statut socio-économique faible sont plus souvent issus d'ethnies minoritaires et ont moins souvent accès aux nouveaux types de traitement telle que la pompe à insuline.

Pour les facteurs génétiques, peu d'articles ont été publiés entre janvier 2011 et juillet 2016, et aucune revue systématique pour la partie génétique n'avait été faite jusque-là. Nous avons donc décidé de ne pas faire de sélection sur l'âge et sur la date de publication des articles. Quatorze articles ont ainsi été sélectionnés parmi les 277 identifiés.

Concernant l'effet des facteurs génétiques, la majorité des études sont des études de gènes-candidats et portent sur de petits échantillons. Ces études ont le plus souvent comme objectif principal d'étudier les complications du DT1 ou d'identifier des gènes associés au DT1. Seules trois études portent sur des cohortes de taille supérieure à 500 individus. A l'exception du gène SLC30A8, aucun gène connu pour être associé au taux d'HbA1c chez des individus non-diabétiques n'a été testé chez les patients souffrant de DT1. Six gènes ont été suggérés par ces études comme étant associés au taux d'HbA1c (Interleukine 6 (IL6) (Ururahy *et al.*, 2015), *Toll-Like Receptor 3* (TLR3) (Assmann, Brondani Lde, Bauer, Canani, & Crispim, 2014), *Tumor Necrosis Factor* (TNF) (Korolija *et al.*, 2010), *Protein Tyrosine Thosphatase, Ton receptor type 22* (PTPN22) (Petrone *et al.*, 2008), *Cytotoxic T-Lymphocyte-Associated protein 4* (CTLA4) (Almasi *et al.*, 2015), *ERB-B2 receptor tyrosine kinase3* (ERBB3) (Kaur *et al.*, 2016)). Une étude combinée de 11 gènes exprimés dans les îlots de Langerhans et régulés par des cytokines a permis de créer un score de risque génétique basé sur le nombre d'allèles à risque pour ces 11 gènes (Brorsson *et al.*, 2016).

Un seul GWAS avait été réalisé, sur la cohorte DCCT/EDIC (EDIC pour *Epidemiology of Diabetes Interventions and Complications*, correspondant à la suite de la cohorte DCCT) (Paterson *et al.*, 2010) composée de patients souffrant de DT1 adultes qui ont été suivis durant 6 ans et demi afin d'étudier l'évolution du DT1 ainsi que la survenue de complications. Dans le cadre de ce GWAS, l'association génétique a été testée dans la cohorte totale ou en stratifiant sur le

traitement (conventionnel qui consiste en au maximum deux injections d'insuline par jour ou intensif pour les patients sous pompe ou avec plus de deux injections). Le SNP le plus significativement associé au taux d'HbA1c est le SNP rs1358030, localisé près du gène *SORCS1*. Le gène *SORCS1*, pour *sortilin-related vacuolar protein sorting 10 domain containing receptor 1*, est associé à différents traits métaboliques chez la souris (Clee *et al.*, 2006; Granhall, Park, Fakhrai-Rad, & Luthman, 2006). De plus, une délétion ou une surexpression de ce gène chez la souris entraîne une sécrétion anormale d'insuline dans les cellules  $\beta$  des îlots de Langerhans (Kebede *et al.*, 2014). Trois autres régions, 14q32.13 (*GSC*, *Goosecoid protein*), 9p22 (*BNC2*, *Basonucline 2*) et 15q21.3 (*WDR72*, *WD repeat-containing protein 72*), ont été identifiées comme étant associées au taux d'HbA1c chez les patients souffrant de DT1, mais n'ont pas été répliquées dans une cohorte indépendante (Paterson *et al.*, 2010).

Notre revue a confirmé que les principaux facteurs associés au taux d'HbA1c sont démographiques, personnels et liés au DT1. Les études génétiques, peu nombreuses, sont souvent peu puissantes et n'ont à ce jour pas été répliquées. Cependant, notre étude souffre de certaines limitations. La plupart des études sur le taux d'HbA1c chez le patient souffrant de DT1 ne sont pas ajustées sur les différents cofacteurs, et notamment sur l'effet du traitement. Un autre problème important est l'hétérogénéité des analyses concernant le taux d'HbA1c, ce qui nous a empêché d'effectuer une méta-analyse.

Un certain nombre d'études concernant les facteurs non génétiques du taux d'HbA1c chez l'enfant et le jeune adulte souffrant de DT1 a été publié après la date du 18 juillet 2016, date à laquelle nous avons fait l'extraction des publications, avec plus de 900 articles qui correspondaient à la recherche avec les mêmes mots clefs. La majorité de ces études retrouve des facteurs identiques que ceux présentés dans la revue de la littérature (voir article ci-après).

Depuis la publication de notre revue, de nombreuses études se sont intéressées à l'impact du statut socioéconomique des patients et de leur famille en lien avec l'ethnie ou l'appartenance à une minorité ethnique (Agarwal *et al.*, 2019; Bergenstal *et al.*, 2017; Mönkemöller *et al.*, 2019; Thomas, Lipsky, Liu, & Nansel, 2018; Zuckerman-Levin *et al.*, 2018). Concernant les facteurs liés au DT1 ou à son traitement, plusieurs études se sont aussi intéressées à l'utilisation des nouvelles technologies ou la mise en place de programme d'éducatifs, et dans les deux cas les effets semblent faibles ou non significatifs (Dłużniak-Gołaska, Panczyk, Szypowska, Sińska, &

Szostak-Węgierek, 2019; Eleftheriadou *et al.*, 2019; Geremia, Fornari, & Tschiedel, 2019; Sun, Malcolm, Wong, Shorr, & Doyle, 2019; Vesco, Jedraszko, Garza, & Weissberg-Benchell, 2018). L'utilisation des pompes à insuline a aussi été étudiée, avec un effet de l'utilisation de la pompe sur le taux d'HbA1c (Blair *et al.*, 2019; Botros, Islam, & Hursh, 2019; Deeb *et al.*, 2019; Karges *et al.*, 2017). L'absence d'effet du type d'insuline a aussi été reporté par plusieurs articles (Bode *et al.*, 2019; Nørgaard, Sukumar, Rafnsson, & Saravanan, 2018; Yamamoto *et al.*, 2019). L'impact de l'implication des parents sur le taux d'HbA1c a aussi été testé dans plusieurs articles et montre que le stress des parents est associé à un taux d'HbA1c plus élevé et que la plus grande implication des parents est à l'inverse associée à un meilleur équilibre glycémique (Ferrito *et al.*, 2019; Healey *et al.*, 2018). Parmi les autres traits étudiés, différentes études se sont notamment intéressées au taux d'HbA1c au diagnostic et à son impact sur l'équilibre glycémique par la suite (Piccini *et al.*, 2019), à l'adhérence des patients au traitement (Haas *et al.*, 2019) et à l'IMC standardisé (Birkebaek *et al.*, 2018). L'ensemble de ces études montre l'intérêt qu'il y a encore à identifier les facteurs influençant le taux d'HbA1c et à caractériser leur effet afin de mieux traiter les patients.

Depuis notre revue, cinq nouvelles études ont cherché à identifier d'autres facteurs génétiques pouvant avoir un rôle dans l'équilibre glycémique chez le patient souffrant de DT1 (Abu El-Ella, Khattab, El-Mekkawy, & El-Shamy, 2018; Rokamp *et al.*, 2018; Slominski *et al.*, 2019; Souza *et al.*, 2018; Syreeni *et al.*, 2019). Quatre de ces études sont des études de gènes-candidats et ont eu pour objectif d'identifier des gènes impliqués dans le développement du DT1 (les gènes CD226 (*Cluster of Differentiation 226*) (Abu El-Ella *et al.*, 2018) et LRP5 (*LDL receptor-related protein 5*) (Souza *et al.*, 2018)) ou associés à des traits liés au DT1 (les gènes ADRB2 (*beta-2 adrenergic receptor*) dans le cadre des hypoglycémies sévères (Rokamp *et al.*, 2018) et CCR5 (*C-C chimiokine de type 5*) avec la dyslipidémie (Slominski *et al.*, 2019)). Deux d'entre elles ne montrent pas d'association significative entre le taux d'HbA1c et les gènes testés (Abu El-Ella *et al.*, 2018; Rokamp *et al.*, 2018). La dernière étude correspond à un GWAS effectué sur une cohorte de patients souffrant de DT1 multicentrique nationale, et a mis en évidence une association entre le taux d'HbA1c et un marqueur près du gène RXFP2 (*Relaxin Family Peptide Receptor 2*) sur le chromosome 13 (Syreeni *et al.*, 2019).

Le contrôle du taux d'HbA1c est complexe, influencé par de nombreux facteurs, qui peuvent être corrélés entre eux. Il semble important, après avoir réalisé cette revue, de les étudier conjointement afin de définir l'effet de chacun d'entre eux ainsi de pouvoir ajuster la dose d'insuline requise par jour (DIR) chez le patient souffrant de DT1, afin de limiter la période de titrage et d'ajustement nécessaire à l'heure actuelle. L'étude conjointe des différents facteurs identifiés a en partie été réalisée pour l'étude préliminaire du projet GENDIR, qui est présentée dans le Chapitre 4.

## 3.2 Article 2 (publié en ligne en septembre 2018)


Received: 16 March 2018 | Revised: 20 July 2018 | Accepted: 23 July 2018

DOI: 10.1002/dmrr.3051

### REVIEW ARTICLE

WILEY

# A systematic review of non-genetic predictors and genetic factors of glycated haemoglobin in type 1 diabetes one year after diagnosis

Emilie Gloaguen<sup>1,2</sup>  | Nathalie Bendelac<sup>3</sup> | Marc Nicolino<sup>3</sup> | Cécile Julier<sup>1,2</sup> | Flavie Mathieu<sup>4</sup>

<sup>1</sup>Inserm UMRS-958, Paris, France<sup>2</sup>Université Paris Diderot, Sorbonne Paris Cité, Paris, France<sup>3</sup>Hôpital Femme-Mère-Enfant, Hospices Civils de Lyon, Bron, France<sup>4</sup>Inserm DISC, siège, Paris, France

#### Correspondence

Emilie Gloaguen, Inserm UMRS-958, Paris, France; Université Paris Diderot, Sorbonne Paris Cité, Paris, France.  
Email: emilie.gloaguen@inserm.fr

#### Present Address

Emilie Gloaguen, Département d'Epidémiologie, Biostatistiques et Recherche Clinique, URC Paris Nord. Hôpital Bichat, Claude-Bernard, Hôpitaux Universitaires Paris Nord Val de Seine, 46 rue Henri Huchard, 75877 Paris Cedex 18, France

#### Funding information

French Ministère de la Recherche; Institut national de la santé et de la recherche médicale

#### Summary

Type 1 diabetes (T1D) results from autoimmune destruction of the pancreatic  $\beta$ cells. Although all T1D patients require daily administration of exogenous insulin, their insulin requirement to achieve good glycaemic control may vary significantly. Glycated haemoglobin (HbA1c) level represents a stable indicator of glycaemic control and is a reliable predictor of long-term complications of T1D. The purpose of this article is to systematically review the role of non-genetic predictors and genetic factors of HbA1c level in T1D patients after the first year of T1D, to exclude the honeymoon period. A total of 1974 articles published since January 2011 were identified and 78 were finally included in the analysis of non-genetic predictors. For genetic factors, a total of 277 articles were identified and 14 were included. The most significantly associated factors with HbA1c level are demographic (age, ethnicity, and socioeconomic status), personal (family characteristics, parental care, psychological traits...) and features related to T1D (duration of T1D, adherence to treatment ...). Only a few studies have searched for genetic factors influencing HbA1c level, most of which focused on candidate genes using classical genetic statistical methods, with generally limited power and incomplete adjustment for confounding factors and multiple testing. Our review shows the complexity of explaining HbA1c level variations, which involves numerous correlated predictors. Overall, our review underlines the lack of studies investigating jointly genetic and non-genetic factors and their interactions to better understand factors influencing glycaemic control for T1D patients.

#### KEYWORDS

children, diabetes mellitus, type 1, glycaemic control, haemoglobin A, glycosylated, predictors

**List of abbreviations:** ACE, Angiotensin I converting enzyme; ADA, Adolescent Diabetes American Diabetes Association; APOE, Apolipoprotein E; BMI, Body mass index; BMI-SDS, Standardized or adjusted BMI scores for age and gender; COBL, Cordon-Bleu WH2 repeat protein; CSII, Continuous subcutaneous insulin infusion or pump; CTLA4, T lymphocyte-associated antigen-4; CTRB1, Chymotrypsinogen B1; DanDiabKids, The Danish registry of childhood and adolescent diabetes; DCCCT, Diabetes Control and Complications Trial; DKA, Diabetic ketoacidosis; DPV, The Diabetes-Patienten-Verlaufsdokumentation or diabetes-patient-documentation; ERBB3, Erb-B2 receptor tyrosine kinase 3; GCK, Glucokinase gene; GSC, Goosecoid homeobox; GSDMB, Gasdermin B; GWAS, Genome-wide association study; HbA1c, Glycated haemoglobin; IL-6, Interleukin-6; IFH1, Interferon induced with helicase C domain 1; IL-10, Interleukin-10; IL-7R, Interleukin-7 receptor; INS, Insulin; ISPAD, International Society for Paediatric; LADA, Latent autoimmune diabetes in adults; MDI, Multiple daily insulin injections; MAGIC, The Meta-Analyses of Glucose and Insulin-related traits Consortium; MeSH, Medical Subject Headings; MODY, Maturity onset diabetes of the young; MTNR1B, Melatonin receptor 1B; PTPN22, Protein tyrosine phosphatase non-receptor type 22; SES, Socioeconomic status; SH2B3, SH2B Adaptor protein 3; SKAP2, Src kinase associated phosphoprotein 2; SNP, Single nucleotide polymorphisms; SORCS1, Sorting 10 domain containing receptor 1; SweDiabKids, The Swedish childhood diabetes registry; T1D, Type 1 diabetes; T1DX, The T1D exchange clinic network; T2D, Type 2 diabetes; TNF- $\alpha$ , Tumour necrosis factor-Alpha; TLR3, Toll-like receptor 3

## 1 | INTRODUCTION

Type 1 diabetes (T1D) is a multifactorial disorder resulting from autoimmune destruction of pancreatic  $\beta$  cells, involving numerous genetic and environmental factors.<sup>1-3</sup> Despite technological advances in T1D treatment and glucose monitoring, glycaemic control remains a day-to-day challenge for T1D patients. Glycated haemoglobin (HbA1c) level represents a stable indicator of long-term glycaemic control, as it reflects the average plasma glucose concentration over approximately 3 months (which corresponds to the average life span of erythrocytes).<sup>4</sup>

Although blood glucose measurement remains the method of choice for T1D diagnosis,<sup>5</sup> HbA1c may also be of interest to exclude or confirm T1D diagnosis,<sup>6,7</sup> with the advantage of its convenience and reproducibility as well as fewer variations due to other factors such as stress or illness.<sup>8</sup> For diabetic patients, good glycaemic control is defined by an HbA1c level lower than 7.5% for children and young adults according to the International Society for Paediatric and Adolescent Diabetes (ISPAD)<sup>9</sup> and the American Diabetes Association (ADA)<sup>10</sup> recommendations. For diabetic adults, the ADA's HbA1c goal is lower than 7%.<sup>11</sup> HbA1c level is also a reliable predictor of the development of long-term complications in T1D patients.<sup>12-14</sup>

A previous review which considered demographic, personal, and interpersonal factors associated with HbA1c level, based on studies published between 1993 and 2010, concluded the effect of ethnicity, socioeconomic status (SES), and family structure on HbA1c level.<sup>15</sup> In a recent review, Mazarello Paes et al reported associations between demographic and T1D-linked factors and HbA1c level in the first year following T1D diagnosis, thus focusing on the honeymoon period.<sup>16</sup>

To our knowledge, no study has been conducted on the heritability of HbA1c level in T1D patients, but several studies exist in non-diabetic and T2D patients, showing an heritability between 55% and 71%.<sup>17-19</sup> Although there is currently no review on genetic factor effects controlling HbA1c level in T1D patients, two reviews<sup>20,21</sup> and a recent genome-wide meta-analysis<sup>22</sup> have reported several loci associated with HbA1c in non-diabetic subjects.

The aim of the present review is to qualitatively synthesize the data from the recent literature concerning both non-genetic predictors and genetic factors of HbA1c level in children and young adult T1D patients, after the first year of the disease.

## 2 | MATERIALS AND METHODS

### 2.1 | Search strategy

Two reviewers (E.G., F.M.) independently searched MEDLINE, Web of Science, and Cochrane Library databases by screening titles, abstracts, and full-texts, and decided about the eligibility of articles. A search was carried out on 18 July 2016. Uncertainties about inclusion/exclusion were resolved by discussion, and only original articles were considered. Since HbA1c is a measure of the metabolic control of diabetes, the selection of non-genetic predictor studies was performed using a combination of free text terms (HbA1c, glycaemic equilibrium, or glycaemic or metabolic control to be found in the title

or abstract) and Thesaurus or MeSH (Medical Subject Headings) terms ("Hemoglobin A, Glycosylated" and "Diabetes Mellitus, Type 1"). Titles and abstracts were double screened to define the inclusion/exclusion status. As we aimed to perform a follow-up of the previous review of Neylon et al published in 2013,<sup>15</sup> we chose a similar search strategy based on the following criteria: (1) publication between January 2011 (studies published before 2011 being included in the previous review<sup>15</sup>) and July 2016; (2) study concerning T1D patients less than 25 years old; and (3) English language.

Only studies with HbA1c level analysis after 1 year of T1D and with sample size greater than 50 participants were included. Studies focusing on mixed populations including type 2 diabetes (T2D), maturity onset diabetes of the young (MODY), latent autoimmune diabetes in adults (LADA), or healthy volunteers were excluded. In the event of duplicate studies, only one was considered. Some studies investigated independent factors on the same large cohort and so have been included in the selection:

- the Danish registry of childhood and adolescent diabetes (DanDiabKids)<sup>23-25</sup>;
- the Diabetes-Patienten-Verlaufsdokumentation or diabetes-patient-documentation (DPV Germany-Austria)<sup>26-30</sup>;
- the T1D exchange clinic network (T1DX)<sup>31-37</sup>;
- the Swedish childhood diabetes registry (SweDiabKids).<sup>38-43</sup>

As no review exists on the genetic factors of HbA1c in T1D patients and due to the low number of studies, a similar search strategy was used, with the additional following inclusion/exclusion criteria: (1) the term "genetic" listed in the key words; (2) no date limit for publications; and (3) a less restrictive age limit (childhood and early adulthood).

A global quality assessment was defined by the following quality criteria: (1) power of the study (based on sample size and the factor effect); (2) statistics: type of statistic test (regression, t-test, ANOVA, correlation, etc.), use of a parametrical or non-parametrical statistic test for limited samples, and verification of the normal distribution of the residuals for regressions; (3) correction for multiple testing and significance threshold; (4) adjustment of HbA1c level on insulin dose requirement; and (5) adjustment on covariates. Studies were then classified as high, intermediate, or low quality based on these quality criteria. Studies with a low quality criterion were excluded.

### 2.2 | Data synthesis strategy

Because of the vast heterogeneity across studies in terms of populations (ascertainment criteria and size), design, follow-up period, and the disparate range of HbA1c level measurements (quantitative or qualitative trait using different cut-offs), we decided to present the results qualitatively. The strength of the association between predictors and HbA1c level was summarized using the Harvest Plot format.<sup>44</sup>

For the non-genetic predictor studies, we extracted the following data: first author, year of publication, study location, sample size, inclusion and exclusion criteria, statistical tests and power, adjustment



of HbA1c level on insulin dose requirement, covariates included in the statistical analysis, correction for multiple testing, clinical features (age, body mass index [BMI], percentage of males in the sample, ethnicity), T1D characteristics (duration, age at onset, type of treatment), mean HbA1c level, and main results with significance thresholds.

For the genetic factor studies, the same data was extracted, with the addition of the following items: specific study design (type of study: candidate gene or genome-wide association study), nucleotide polymorphisms (SNP) considered, associated SNPs, Hardy-Weinberg equilibrium statistics, and association test *P*-value.

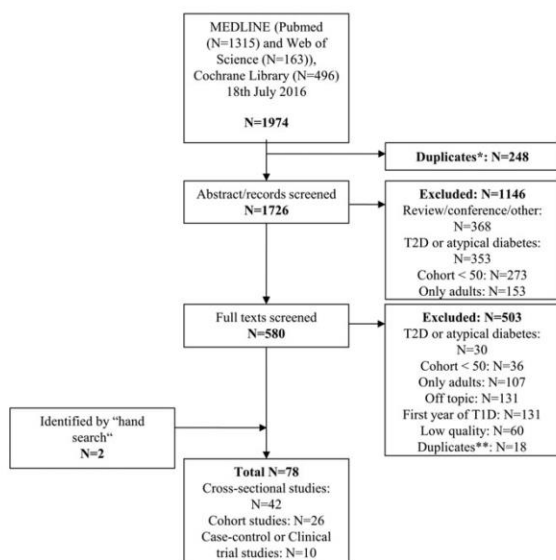
Genes were classified into four groups using the EBI GWAS catalogue<sup>45</sup> (<https://www.ebi.ac.uk/gwas/>) and Gene Ontology (string)<sup>46</sup> (<https://string-db.org>). The following groups were considered: "Immunity" for genes associated or implicated in immune or inflammation process; "Autoimmune-disease" for genes associated or implicated in autoimmune-disease; "Metabolism" for genes associated or implicated in metabolic trait or process (glucose or lipid); and "Erythrocyte" for genes associated or implicated in red cell life. Genes associated with HbA1c level but not with one of the four previous groups were classified as "Others".

### 3 | RESULTS

#### 3.1 | Review of non-genetic predictor studies

##### 3.1.1 | Review process and quality appraisal

A total of 1974 articles were identified and 78 were finally retained by the review procedure (Figure 1). Twenty-six studies had a cohort design, 42 were cross-sectional, and the rest were case/control studies (Figure 1 and supplementary Tables S1, S2, S3, and S4).



**FIGURE 1** Flow diagram of literature search and selection for non-genetic factors studies\* articles duplicated electronically; \*\* cohort presented at different times of the follow-up

Association results between predictors and HbA1c level are summarized in Figure 2.

In total, 38 studies provided high quality level,<sup>24,25,27-30,32-37,39,40,43,47-69</sup> and 40 studies provided intermediate level.<sup>41,42,70-107</sup>

The sample size range is 70 to 30 708 patients with a median sample size of 210 T1D patients. Thirty-two studies were performed in several European countries,<sup>24,25,27-30,42,43,47,49,50,52,57,58,61-63,66,72,73,75,82,94,107</sup> 31 in USA,<sup>32-37,48,56,59,60,64,65,67,71,76,80,81,84-87,89,90,92,95,98,99,102,103,105,106</sup> one in Canada,<sup>51</sup> three in Israel,<sup>53,100,101</sup> and the rest in other countries (one in Japan,<sup>88</sup> one in New-Zealand,<sup>78</sup> three in Turkey,<sup>70,79,83</sup> one in Australia,<sup>69</sup> and three are multinational<sup>55,77,93</sup>). Most studies (54 studies) were performed on T1D children (median of mean age reported in articles 12.5 [range: 4.1-17.7]), 22 included both children and young adults (median of mean age of 14.8 [range: 10.1-21.6]), and one included only young adults (mean age of 19.4 ± 0.9 [range: 18-21]). For details, see supplementary Table S4.

#### 3.2 | Demographic predictors

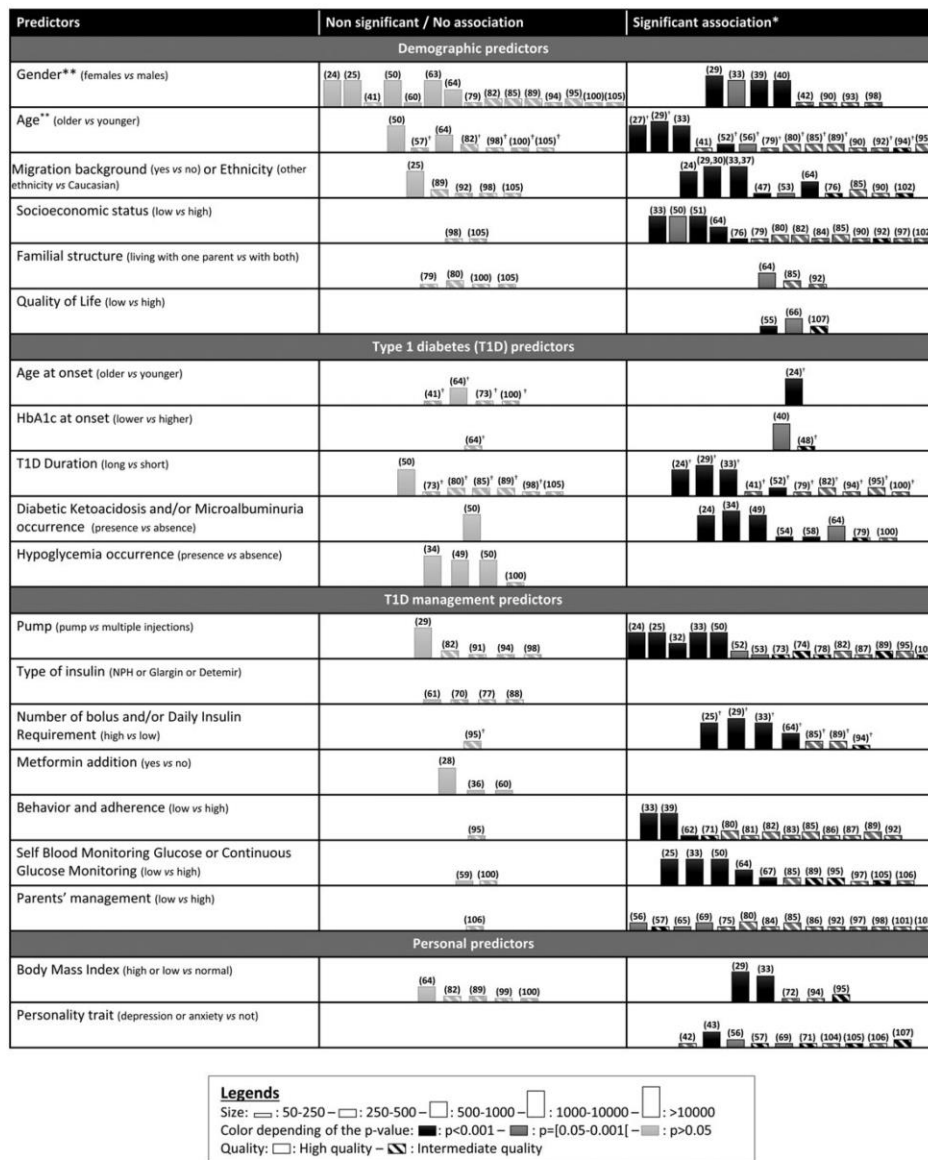
Age, gender, ethnicity, and SES are usually considered as demographic characteristics that may affect HbA1c level. Although only a few studies have directly addressed the influence of age and gender<sup>27,35,40</sup> on HbA1c level, these two factors were investigated in several articles, where they have been considered as confounding factors in the analysis, based on clinical considerations.

Gender effect was still controversial: eight studies, with variable sample sizes (N = 115 to 30 708), reported a significant gender effect, with girls showing increased HbA1c level compared with boys,<sup>29,33,39,40,42,90,93,98</sup> whereas 15 studies showed no association, nine of which had intermediate quality,<sup>41,79,82,85,89,94,95,100,105</sup> and six had high quality,<sup>24,25,50,60,63,64</sup> based on our criteria.

In most studies, age was significantly and positively correlated with HbA1c level.<sup>27,29,33,41,52,56,79,80,82,85,89,90,92,94,95</sup> Four small-scale studies<sup>57,82,98,105</sup> and one performed in a very early age at T1D onset<sup>100</sup> reported non-significant association. Age effect is not linear<sup>29,40,50,90,100</sup>; a slight decrease of HbA1c level was observed between the age of 2 and 7, an increase between 7 and 19, and a decrease between 19 and 25. This non-linear effect was confirmed in an extended cohort study.<sup>108</sup> Although the age effect remained significant after adjusting for daily insulin dose injection,<sup>94</sup> it was found to be restricted to patients treated by injections, but not by pump.<sup>25</sup>

Gender and age are not independent: no gender effect was observed before the age of 13<sup>41,50,63,95,100</sup> and in young adults,<sup>35</sup> contrasting with an increase of HbA1c level in girls compared with boys between 13 and 21 years old.<sup>40,42</sup> This age-gender interaction may be related to puberty: pubertal patients having higher HbA1c levels than pre-pubertal ones and girls having higher HbA1c levels than boys during puberty only.<sup>27,79</sup>

Ethnicity and migration background were associated with HbA1c level in Europe,<sup>24,25,29,30,47</sup> America,<sup>37,64,76,85,89,90,92,98,102,105</sup> and Israel.<sup>53</sup> In the US, HbA1c level was found to be higher in T1D patients belonging to ethnic minority groups, with a mean HbA1c level difference varying between 0.9% (9.8 mmol/mol) and 1.5% (16.4 mmol/mol) in African-Americans compared with Caucasian-



**FIGURE 2** Predictors of higher HbA1c level. \*: association reported in the referenced articles (shown in brackets); \*\*: articles focusing on the age and gender correlation are not presented in this figure; †: factors were studied as a continuous variable or with as or more than three categories

Americans.<sup>37,64,76,85,90,102</sup> Hispanic ethnicity effect compared with Caucasian patients is discordant across studies, some showing no difference<sup>37,64,102</sup> while another reported higher HbA1c levels in Hispanic<sup>90</sup> patients. In Israel, Bedouin patients have a higher HbA1c level than Jewish<sup>53</sup> patients. In Germany, patients who have a Turkish mother have a higher HbA1c of 0.3% (3.3 mmol/mol) compared with patients with a German or Austrian mother.<sup>30</sup> Children with both parents born in the same country showed lower HbA1c levels (1.1% [12.0 mmol/mol]) than children having one parent born outside the country.<sup>47</sup> Association between migration and HbA1c level may be an indirect consequence of other factors related to the migrant condition, such as access to specific treatments including pump<sup>25</sup> or SES.

HbA1c level and SES have been shown to be correlated.<sup>33,37,50,51,64,76,79,80,82,84,85,90,92,97,102</sup> In most studies, SES is defined by the parental education level, household income, type of insurance and/or type of neighbourhood, or a specific index based on these factors.<sup>109</sup> A higher HbA1c level is reported for patients having either a low household income<sup>33,51,64,80</sup> or a low health insurance.<sup>33,64,90</sup> A low HbA1c level is associated with parents' high education level,<sup>33,64,84</sup> regardless of which parent.<sup>79</sup> However, the employment status of the mother (unemployed versus working full or part time) was reported to have no effect on HbA1c level.<sup>79</sup>

As discussed above, the SES effect may interact with other factors. In the United States, where Caucasian-American patients frequently



reported higher SES status than African-Americans,<sup>37,76</sup> differences in HbA1c levels between these populations remained significant after taking into account the SES status, thus suggesting independent contributions of these two factors.<sup>37,64</sup> As expected, a lower quality of life was associated with a higher HbA1c level,<sup>55,66,107</sup> even when taking into account confounders such as the depressive syndrome.<sup>107</sup>

### 3.3 | Type 1 diabetes-related predictors

The impact of parameters related to T1D onset on HbA1c level after the first year of diabetes has been poorly investigated to date and has been limited to three parameters: HbA1c level,<sup>40,48,64</sup> age,<sup>24,41,64,73,100</sup> and occurrence of diabetic ketoacidosis (DKA).<sup>64,110</sup> HbA1c level at onset is consistently correlated to HbA1c level after 2 or 3 years of T1D.<sup>40,48,64</sup> Studies investigating age at onset were discordant, one showed a significant effect on HbA1c level (increase of HbA1c level with age),<sup>24</sup> and four were non-significant, possibly as a consequence of study design: one study ascertained patients through a very early age at onset (lower than 6.5 years<sup>100</sup>) and another considered the mean HbA1c level over the 8 years of the study duration.<sup>73</sup> The presence of DKA at onset was associated with an increase of HbA1c level of 0.3% (3.3 mmol/L), and this finding remained significant after adjusting for age and treatment.<sup>24,64</sup>

A shorter T1D duration was significantly associated with a lower HbA1c level,<sup>24,29,33,41,52,79,82,94,95,100</sup> even after adjusting for age and treatment.<sup>29,33</sup> Several articles reported a strong association between the occurrence of DKA and/or microalbuminuria and HbA1c level,<sup>34,49,54,58,79,100</sup> although no association between HbA1c level and hypoglycaemia was observed.<sup>34,49,50,100</sup>

### 3.4 | T1D management-related predictors

Currently, children and young adults with T1D have better glycaemic control when they are treated by intensive insulin therapy (namely multiple daily insulin injections [MDI]) or pump (namely continuous subcutaneous insulin infusion [CSII]) than by conventional therapy (one to three injections).<sup>29,85</sup> Over the last decades, pump therapy has been used increasingly in children and adolescents, and HbA1c level was significantly lower in patients treated by CSII compared with MDI,<sup>24,25,32,33,52,53,73,74,78,80,87,89,95,105</sup> with the exception of four small sample studies.<sup>82,91,94,98</sup> This difference was observed for each class of age over the first 5 years of treatment and remains significant after adjusting for duration<sup>52</sup> and other covariates.<sup>33</sup> Patients treated initially with CSII and then switched to MDI reported greater HbA1c level than patients using only CSII therapy.<sup>105</sup> HbA1c level also decreased according to the duration between T1D onset and the beginning of CSII treatment.<sup>24,73</sup> Regardless of the disease treatment, studies consistently reported an increase of daily insulin dose requirement during the first 5 years of the disease.<sup>25,29,33,64,85,89,94</sup> In most studies, the number of bolus was negatively correlated to HbA1c level.<sup>25,33</sup>

Multiple daily insulin injections insulin regimen involves long- and rapid-acting insulin analogues. All long-acting insulin analogues (glargine, detemir, or NPH) have similar effects on HbA1c level,<sup>50,61,63,70,77,88</sup> although patients under detemir require higher insulin doses than patients under NPH<sup>61</sup> or glargine.<sup>70</sup> A recent review

reported HbA1c level improvement over the last decades for both MDI and CSII insulin regimen with insulin lispro analogue.<sup>111</sup> The effect of premixed insulin regimen (composed of different percentage of long or intermediate and rapid acting insulin) on HbA1c level compared with basal/bolus and pump therapy is still controversial.<sup>50,63</sup>

The addition of metformin to the insulin treatment has been proposed for overweight T1D adolescents in order to improve glycaemic control. However, after adding metformin for 6 months, HbA1c level remains unchanged,<sup>28,36,60</sup> while gastrointestinal adverse events may occur.<sup>36</sup>

Good adherence and optimal management are critical to achieve good glycaemic control. Patients with good diabetes management, evaluated by the Self-Care Inventory,<sup>71,80,81,87,112</sup> the Diabetes Behavior Rating Scale,<sup>85,113</sup> the Diabetes Self-Management Profile,<sup>86,114</sup> the Diabetes Interview,<sup>92,115</sup> or a composite score,<sup>33</sup> have lower HbA1c levels.<sup>33,71,80,81,85-87,92</sup> In contrast, missing injections<sup>33</sup> or poor self-monitoring of blood glucose<sup>25,33,50,64,67,85,89,95,98,105,106</sup> was correlated with increased HbA1c level.

To optimize prandial insulin injection, patients may use automatic bolus calculator and/or carbohydrate counting. This practice is associated with a decrease of HbA1c level (0.5% [5.5 mmol/mol] after 12 months<sup>62,83</sup> and 0.9% [9.8 mmol/mol] after 18 months<sup>83</sup>). Surprisingly, participation in a structured education course (using interactive and practical learning activities about carbohydrate counting and insulin dose adjustment in everyday life) failed to improve HbA1c level after 24 months.<sup>96</sup> Only one study tested the impact of continuous glucose monitoring on HbA1c level and was inconclusive.<sup>59</sup> Adherence to dietary and physical activity guidelines reduces HbA1c level.<sup>39</sup> In adolescents, daily media consumption time (time spent watching television and using computers),<sup>82</sup> lower fibre intake, and higher fat intake<sup>89</sup> are associated with increased HbA1c level.

In paediatric T1D patients, parental involvement in T1D management is crucial,<sup>116</sup> albeit complex and challenging, particularly during adolescence. Parental involvement may be direct (eg, administering insulin injections) or indirect (eg, remembering medical appointments). In most families, mothers are primary caregivers whereas fathers are often involved only in the case of poor glycaemic control.<sup>56,86</sup> For young patients, direct parental involvement is associated with better HbA1c level. By contrast, for teenagers, parents' direct involvement in daily T1D management decreases the patient's ability to manage T1D and leads to higher HbA1c level.<sup>116</sup> In adolescents, authoritative non supportive management is associated with better glycaemic control.<sup>101</sup> Good parental knowledge and perception of T1D improve glycaemic control,<sup>68</sup> with a decrease of HbA1c level of 0.5% (5.0 mmol/mol) when parents are able to count carbohydrate intake and of 0.4% (3.9 mmol/mol) when parents have a good perception of self-efficacy with regards to diabetes.<sup>97</sup> Mother-father diabetes management discrepancies lead to worse glycaemic control.<sup>65</sup> Surprisingly, a study reported that having a family history of diabetes was associated with poorer HbA1c level.<sup>24</sup>

### 3.5 | Personal predictors

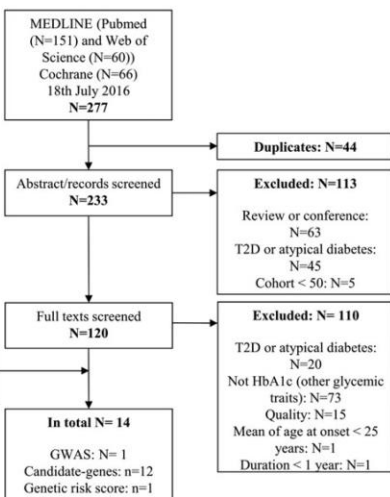
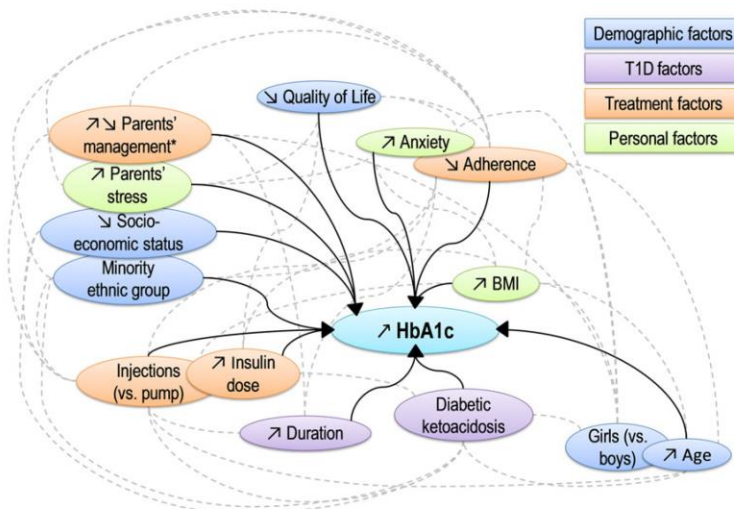
Standardized or adjusted BMI scores for age and gender (BMI-SDS)<sup>94,99,117</sup> were positively correlated with HbA1c level, a

higher proportion of patients with HbA1c greater than 9% being in the elevated BMI group.<sup>29,33</sup> Surprisingly, patients with the lowest BMI-SDS also reported higher HbA1c level.<sup>29</sup> Moreover, an increase of HbA1c levels during a 5-year follow-up period was correlated to BMI-SDS variation.<sup>72,95</sup> These results may be age-dependent since no effect was observed for patients less than 10 years old.<sup>100</sup> Patients whose parents were of low or normal weight have lower HbA1c level than patients with overweight parents, probably reflecting a family environmental effect.<sup>99</sup>

The impact of personality traits on HbA1c level, such as anxiety, depression, and distress, has been recently reviewed.<sup>118,119</sup> These traits are all associated with higher HbA1c levels. In our present review, five recent studies on depression<sup>71,105-107</sup> and its associated traits,<sup>42,43</sup> such as lethargy or eating disorders, were selected and reported an increase of HbA1c level in patients. Anxiety and depression may interact with other factors like gender or T1D duration.<sup>119</sup> Other psychological traits (emotional processing, distress<sup>57</sup> or low conscientiousness and/or agreeableness<sup>69</sup>) were associated with an increase of HbA1c level, as well as memory or learning disabilities.<sup>41,56</sup> Diabetes acceptance and illness perception were associated with good T1D management, leading to lower HbA1c level,<sup>81</sup> with a more important contribution in females, explaining 23% of the variance of HbA1c level.<sup>104</sup>

The family structure can be defined by the parental marital status<sup>79,85,100,105</sup> or by living with one or both parents.<sup>64,85,92,98</sup> The family structure effect on HbA1c level is controversial, possibly reflecting the lack of power of small sample studies.<sup>79,80,100,105</sup> Living with both parents<sup>64,85,92</sup> improves HbA1c level. A high level of parental-reported family conflict is associated with a worse evolution of HbA1c level, as it impacts the quality of T1D management and adherence as well as increases the child's level of stress.<sup>85,98</sup> Parental general stress and parental distress specifically linked to T1D are associated with an increase of HbA1c level.<sup>57,84</sup> Parental depressive symptoms lead to their child's increased HbA1c level (around 0.2% [2.2 mmol/mol]),<sup>75,80</sup> whether it is the mother<sup>75,92,103</sup> or the father<sup>80</sup> who is concerned.

To conclude, several non-genetic predictors are significantly associated with HbA1c level and most of them interact (Figure 3), which demonstrates the importance of studying them conjointly.



### 3.6 | Genetic factors

Heritability of HbA1c level has never been estimated in T1D patients but has been estimated in non-diabetic Caucasian individuals between 55% and 71%,<sup>17-19</sup> suggesting the contribution of genetic factors.

A total of 277 genetic studies have been conducted in T1D patients, and 14 articles were retained by the review procedure (Figure 4). Results are provided in Table 1 and in Figure 5 (Table S5).

### 3.7 | Candidate gene studies

All genetic studies except one have used a candidate gene approach. The impact of the candidate genes on HbA1c level was rarely the principal objective of the selected studies, most of them investigating HbA1c level as part of a genetic study or complications of T1D (Table 1 and Figure 5). Most association studies were performed on small sample sizes, without adjusting for insulin dose requirement or

FIGURE 4 Flow diagram of literature search and study selection for genetic factors

FIGURE 3 Schematic representation of factors associated with an increase of HbA1c level reported in selected papers:black arrow: factors associated with HbA1c level; dotted grey line: significant interactions between HbA1c factors. \*: Parents' authoritative management effects depend on the patient's age



**TABLE 1** Candidate gene studies of HbA1c levels

Reference	N	Age (yr; Mean ± SD)	Age at Onset (yr; Mean ± SD)	Region	Gene	Classification	SNP	Studied Trait	P-Value*
Petrone et al 2008 <sup>120d</sup>	120	15.9 ± 7.8	14.9 ± 7.8	1p13.2	PTPN22	Autoimmune-disease; immunity	rs2476601	HbA1c	P = 0.01
Nielsen et al 2011 <sup>121e</sup>	257	10.1 ± 3.7	9.1 ± 3.7	1p13.2	PTPN22	Autoimmune-disease; immunity	rs2476601	IDAAC HbA1c	NP NP
Mosaad et al 2011 <sup>122</sup>	104	12.0 ± 3.4	8.22 ± 3.26	2q33.2	CTLA4	Autoimmune-disease; immunity	rs231775	HbA1c	P = 0.33
Almasi et al 2015 <sup>123</sup>	153	23.6 ± 9.1	13.8 ± 7.5	2q33.2	CTLA4	Autoimmune-disease; immunity	rs5742909	HbA1c	P < 0.001
Assmann et al 2014 <sup>124</sup>	449	33.5 ± 13.3	16	4q35.1	TLR3	Immunity	rs11721827 rs5743313 rs13126816 rs7668666 rs3775291	HbA1c	P = 0.013 P = 0.0001 P = 0.22 P = 0.093 P = 0.954
Spoletini et al 2007 <sup>125d</sup>	120	15.9 ± 7.8	14.9 ± 7.8	6p21.32	HLA	Autoimmune-disease; immunity; metabolism	Haplotype <sup>a</sup>	HbA1c	NP
Korolija et al 2010 <sup>126</sup>	170	37	22	6p31.33	TNF	Autoimmune-disease; immunity	rs361525 rs1800629	HbA1c	P = 0.022 P < 0.05
Ururahy et al 2015 <sup>127</sup>	144	11.9 ± 4.0	6.7 ± 3.5	7p15.3	IL6	Immunity	rs1800795 rs1800795	HbA1c	P = 0.029 P = 0.703
Kaur et al 2016 <sup>128</sup>	1133	8.0 ± 2	<11	12q13.2	ERBB3	Autoimmune-disease; metabolism; immunity	rs2292239	IDAAC HbA1c	P = 0.007 P = 0.019
Bouhanick et al 1999 <sup>129</sup>	76	16 ± 3	9 ± 4	17q23.3	ACE	Erythrocyte	InDel	HbA1c	NP
Jacobson et al 2009 <sup>130f</sup>	1093	27	NP	17q23.3	ACE	Erythrocyte	rs4340	HbA1c	NP
Johannesen et al 2011 <sup>131</sup>	1037	10.0 ± 3.8	<15	17q23.3	ACE	Erythrocyte	InDel	HbA1c	NP
Jacobson et al 2009 <sup>130f</sup>	1093	27	NP	19q13.32	APOE	Metabolism	Haplotype <sup>b</sup>	HbA1c	NP
Brorsson et al 2016 <sup>132e</sup>	257	10.1 ± 3.7	9.1 ± 3.7	Number risk alleles for 11 cytokine induced genes <sup>c</sup>				IDAAC	P < 0.01

\*P-value reported in the article. NP = Not provide.

<sup>a</sup>High risk: DRB1\*03-DQB1\*0201/DRB1\*04-DQB1\*0302 (DRB1\*04 different from 0403); moderate risk: DRB1\*04-DQB1\*0302/DRB1\*04-DQB1\*0302, DRB1\*03-DQB1\*0201/DRB1\*03-DQB1\*0201, DRB1\*04-DQB1\*0302/X and DRB1\*03/X (X different from DRB1\*03, DRB1\*04-DQB1\*0302 or DQB1\*0602/03); low risk: all other genotypes.

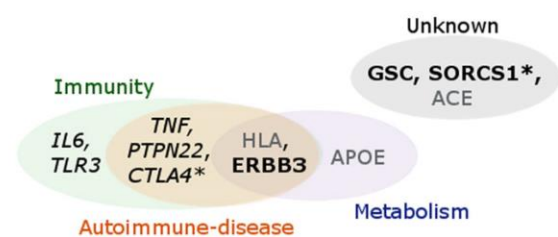
<sup>b</sup>APOE: rs7412 (e2/e3) and rs429358 (e3/e4).

<sup>c</sup>Corresponding to SH2B3; IFIH1; TNFAIP3; GSDMB; IL10; IL7R; CTRB1; CTSH; COBL; SKAP2; INS. IDAAC, insulin dose adjusted HbA1c; NS, non-significant; NP, exact P-value are not provided in the original paper, but the association was indicated as not significant or P > 0.05.

<sup>d</sup>Same cohort of Italian patients.

<sup>e</sup>Same cohort of European and Japanese patients.

<sup>f</sup>Data from DCCT study.



**FIGURE 5** Genes tested to influence HbA1c level in T1D. **Bold:** significant association after adjusting for covariates and multiple corrections. *Italic:* significant association reported in the original article, but not significant after adjusting for covariates and/or for multiple testing. Grey: non-significant association. Genes are grouped according to their reported disease association or known function.\*: Unknown linked with diabetes

other non-genetic predictors of HbA1c level, and should therefore be interpreted with caution. These candidate gene association studies have never been or have been poorly replicated, except for angiotensin I converting enzyme (ACE) gene.<sup>129-131</sup> Five of the 13 selected studies reported significant or suggestive associations, which were however never replicated.<sup>120,123,124,126,128</sup>

The most robust association was reported with the Erb-B2 Receptor Tyrosine Kinase 3 (ERBB3) rs2292239 T1D risk SNP and HbA1c level. The association remains significant after adjusting for insulin dose, gender, age at onset, and diabetes duration,<sup>128</sup> but this finding has not been replicated yet. The Protein Tyrosine Phosphatase Non-receptor type 22 (PTPN22) T1D risk allele was associated with HbA1c level,<sup>120</sup> but this finding was not replicated in an independent study that used the insulin dose-adjusted HbA1c measure (IDAAC), defined by Mortensen et al in 2009.<sup>121,133</sup> Associations reported for the

candidate genes interleukin-6 (IL-6),<sup>127</sup> T lymphocyte-associated anti-gen-4 (CTLA4),<sup>122,123</sup> and tumour necrosis factor-Alpha (TNF- $\alpha$ )<sup>126</sup> genes did not remain significant after correcting for multiple testing. Another study reported suggestive TLR3 association with HbA1c level, but these results are inconclusive considering that a difference of age at onset of 7 years was noted between the two studied groups and so was possibly responsible for the effect.<sup>124</sup>

Three other *loci* were tested: (1) HLA-DRB1 and -DQB1 haplotypes,<sup>126</sup> conferring the highest risk for T1D<sup>125</sup>; (2) the ACE,<sup>129-131</sup> for its possible role in diabetic nephropathy and hypoglycaemia and (3) apolipoprotein E (APOE).<sup>130</sup> None of these genes was significantly associated with HbA1c in T1D patients. Lack of association between ACE polymorphism and HbA1c level was reported in three independent robust studies.<sup>129-131</sup>

A recent study investigated a genetic risk score of 11 T1D risk variants in islet-expressed and cytokines-regulated genes (COBL, CTRB1, GSDMB, IFH1, INS, IL10, IL7R, SH2B3, SKAP2): HbA1c level, adjusted or not for insulin dose and for age, gender, and HLA risk groups, was increased by 0.2% per risk allele after 1 year of T1D.<sup>132</sup>

### 3.8 | Whole genome association study

To date, only one GWAS has been carried out to search for loci underlying genetic control of HbA1c in T1D.<sup>134</sup> Genetic association was investigated on repeated measures of HbA1c level over a mean of 6.5 years in 1304 T1D patients from the Diabetes Control and Complications Trial (DCCT) cohort and then carried forward for replication in the MAGIC sample of non-diabetic subjects.<sup>135</sup> Only one SNP (rs1358030) located near the sorting 10 domain containing receptor 1 (SORCS1) gene achieved the genome-wide significant threshold ( $P = 3.8 \times 10^{-10}$ ). SORCS1 is known to be associated with metabolic traits in mice,<sup>136,137</sup> and deletion or over-expression of this gene in mice leads to an abnormal insulin secretion in  $\beta$  cells.<sup>138</sup> A second SNP (rs111624318), located near the Gooseoid homeobox (GSC) gene was suggestive ( $P = 2.7 \times 10^{-7}$ ). Both SNPs failed to be replicated in the MAGIC sample,<sup>134</sup> and these results have not been replicated yet.

## 4 | DISCUSSION

We systematically screened articles published between January 2011 and July 2016, focusing on predictors of HbA1c level in T1D. Our review mostly agrees with the previous review from Neylon et al<sup>15</sup> published in 2013. Our review concludes that poorer glycaemic control (ie, elevated HbA1c level) is associated with demographic and personal factors including ethnicity, SES, family structure and family conflict, personal characteristics such as anxiety and depressive symptoms, diabetes-related family stress, conflict, and stress as well as comorbidities such as eating disorders. Our review mainly agrees with the previous conclusions from Neylon et al regarding the demographic factors, with the possible exception of gender that was reported to have no effect on HbA1c level,<sup>15</sup> while our review supports a gender effect, which is likely secondary to puberty rather than a primary effect.

Our review also highlights the better glycaemic control achieved by T1D patients treated by insulin pump therapy compared with patients treated by multi-injections<sup>25,32,52,53,74,78,95</sup>; disagreeing with a previous review<sup>139</sup> was however based on two small scale studies (less than 40 patients). In our present review, new factors have been reviewed, such as T1D onset parameters, duration, and complications of T1D.

Recently, Paes et al published a systematic review article on factors affecting HbA1c level during the first year of T1D, therefore focusing on the honeymoon period.<sup>16</sup> Interestingly, demographic predictors associated with HbA1c level during the first years of T1D, including older age, female gender, and minority group, are also associated with HbA1c level after the first year of diabetes, as shown in our review. Therapy and management predictors were also found to be similar in both periods. However, some predictors investigated especially during the first year, including C-peptide or antibody positivity, are specifically tested at the beginning of T1D and thus were not included in our review. Nevertheless, a recent paper investigated the link between HbA1c level and C-peptide in a cohort of T1D patients with a duration of illness between 1 and 5 years. This paper reported a significant association between HbA1c level and C-peptide, with a lower HbA1c level for patients with higher amounts of secreted C-peptide.<sup>139</sup> This result is consistent with the conclusions presented in the systematic review focusing on the honeymoon period.<sup>16</sup>

Robust predictors of HbA1c level in T1D are puberty, ethnicity, SES, psychological dimensions, type of treatment, and management of T1D. Several arguments are in favour of their interactions. For example, direct parental involvement improved HbA1c level during childhood, whereas it was associated with a higher HbA1c level during puberty. This interaction may be explained by correlated factors such as hormonal or psychological changes.<sup>140,141</sup> Similarly, ethnicity and SES may affect the access to pump therapy, thereby impacting HbA1c level, especially for young Caucasian girls.<sup>37,78</sup> Finding the primary causative factor or the respective contribution of correlated factors remains a challenge. More complex interactions between factors may also exist, including between BMI and age at onset, diabetes duration, gender, insulin dose or regimen,<sup>142</sup> and between gender and insulin dose.<sup>143</sup>

Genetic factors have been poorly studied and usually performed using candidate gene studies in small samples despite the widening of the selection criteria. In the absence of previous reviews of genetic factors controlling HbA1c level in T1D, we chose to select all articles reporting a genetic study of HbA1c in T1D patients with disease onset during childhood or early adulthood. Most studies examined a unique candidate gene, without adjusting for treatment or non-genetic predictors and are poorly replicated. Some studies concluded to the absence of association, but were clearly underpowered, and other studies reported a "significant" association which may be attributable to confounding factors or did not remain significant when applying Bonferroni correction for multiple testing. The most robust association is reported for a SNP located near the ERBB3 gene,<sup>128</sup> which however remains to be replicated. The two genes (SORCS1 and GSC) detected by GWAS were not replicated in 1486 T2D patients as part of the Norwegian population-based cohort (HUNT2).<sup>144</sup> Several meta-



analysis or GWAS identified more than 60 SNPs associated with HbA1c level in non-diabetic individuals in different cohorts from European, African, East Asian, and South Asian ancestry<sup>22,135,145-152</sup> but have never been investigated in T1D patients. These loci were classified in three functional groups: glycaemic, erythrocytic, or unclassified function. Moreover, a recent replication study on the general Greenlandic Inuit population of 15 HbA1c associated loci previously reported in non-diabetic individuals confirmed only two of these loci, the glucokinase gene (GCK) and the melatonin receptor 1B (MTNR1B) in this more homogeneous population.<sup>153</sup>

Our review featured several limitations. We used stringent criteria to select papers and define a specific global quality assessment, possibly leading to restricting the number of eligible studies. Moreover, because of heterogeneity across studies (designs, populations, sample size, demographic characteristics ...) <sup>25,29,50,64,93,154-156</sup> and as measures of HbA1c level are quantitative or binary or defined by the mean of repeated measures, and adjusted or not—for treatment and confounder effects, only a qualitative study has been possible. Indeed, as the overall quality of the selected studies was intermediate to high, our overall results are reliable. Moreover, HbA1c level may also reflect erythrocytic function, which was not systematically addressed by our review strategy and therefore may be under-represented. Our review does not allow to judge to judge whether a particular factor is an independent predictor of HbA1c or whether it predicts the average glycaemia, as these traits are highly correlated. Moreover, some genetic factors could be specifically involved in the honeymoon period,<sup>157</sup> which was excluded in our study.

We can summarize that studies investigating jointly genetic and non-genetic factors are still lacking. Further studies taking into account their correlations and interactions are needed in order to understand factors influencing glycaemic control for T1D patients.

#### ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

#### COMPETING INTERESTS

The authors have declared that no competing interests exist.

#### FUNDING

This work was supported by the Institut National de la Santé et de la Recherche Médicale (Inserm), and Emilie Gloaguen was supported by the French Ministère de la Recherche (Bourse doctorale 2014-2017).

#### AUTHORS' CONTRIBUTIONS

Two reviewers E.G. and F.M. independently searched MEDLINE, Web of Science, and Cochrane Library databases, by screening titles, abstracts, and full-texts and decided about the eligibility of articles. Uncertainties about inclusion/exclusion were resolved by discussion. E.G. and F.M. analysed the data and wrote the manuscript. All authors discussed the results and implications and revised and commented the manuscript at all stages.

#### ORCID

Emilie Gloaguen  <http://orcid.org/0000-0002-5205-2262>

#### REFERENCES

- Pociot F, Lernmark A. Genetic risk factors for type 1 diabetes. *Lancet*. 2016;387(10035):2331-2339.
- Rewers M, Ludvigsson J. Environmental risk factors for type 1 diabetes. *Lancet*. 2016;387(10035):2340-2348.
- Todd JA. Etiology of type 1 diabetes. *Immunity*. 2010;32(4):457-467.
- Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c test in diagnosis and prognosis of diabetic patients. *Biomark Insights*. 2016;11:95-104.
- Association AD. 2. Classification and diagnosis of diabetes. *Diabetes Care*. 2017;40(Supplement 1):S11-S24.
- Ehehalt S, Gauger N, Blumenstock G, et al. Hemoglobin A1c is a reliable criterion for diagnosing type 1 diabetes in childhood and adolescence. *Pediatr Diabetes*. 2010;11(7):446-449.
- Committee IE. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care*. 2009;32(7):1327-1334.
- Herman WH, Fajans SS. Hemoglobin A1c for the diagnosis of diabetes: practical considerations. *Pol Arch Med Wewn*. 2010;120(1-2):37-40.
- Rewers MJ, Pillay K, de Beaufort C, et al. ISPAD Clinical Practice Consensus Guidelines 2014. Assessment and monitoring of glycemic control in children and adolescents with diabetes. *Pediatr Diabetes*. 2014;15(Suppl 20):102-114.
- ADA. 12. Children and Adolescents: Standards of Medical Care in Diabetes-2018. *Diabetes Care*. 2018;41(Suppl 1):S126-s136.
- ADA. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2018. *Diabetes Care*. 2018;41(Suppl 1):S55-s64.
- Chalew SA, McCarter RJ, Hempe JM. Biological variation and hemoglobin A1c: relevance to diabetes management and complications. *Pediatr Diabetes*. 2013;14(6):391-398.
- Nordwall M, Abrahamsson M, Dhir M, Fredrikson M, Ludvigsson J, Arnqvist HJ. Impact of HbA1c, followed from onset of type 1 diabetes, on the development of severe retinopathy and nephropathy: the VISS Study (Vascular Diabetic Complications in Southeast Sweden). *Diabetes Care*. 2015;38(2):308-315.
- Nordwall M, Arnqvist HJ, Bojestig M, Ludvigsson J. Good glycemic control remains crucial in prevention of late diabetic complications—the Linköping Diabetes Complications Study. *Pediatr Diabetes*. 2009;10(3):168-176.
- Neylon OM, O'Connell MA, Skinner TC, Cameron FJ. Demographic and personal factors associated with metabolic control and self-care in youth with type 1 diabetes: a systematic review. *Diabetes Metab Res Rev*. 2013;29(4):257-272.
- Mazarello Paes V, Charalampopoulos D, Edge J, Taylor-Robinson D, Stephenson T, Amin R. Predictors of glycemic control in the first year of diagnosis of childhood onset type 1 diabetes: a systematic review of quantitative evidence. *Pediatr Diabetes*. 2017.
- Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA. A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. *Diabetes*. 2002;51(3):833-840.
- Mills GW, Avery PJ, McCarthy MI, et al. Heritability estimates for beta cell function and features of the insulin resistance syndrome in UK families with an increased susceptibility to type 2 diabetes. *Diabetologia*. 2004;47(4):732-738.
- Pilia G, Chen WM, Scuteri A, et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet*. 2006;2(8):e132.
- Marullo L, El-Sayed Moustafa JS, Prokopenko I. Insights into the genetic susceptibility to type 2 diabetes from genome-wide association studies of glycaemic traits. *Curr Diab Rep*. 2014;14(11):551.

21. Soranzo N. Genetic determinants of variability in glycosylated hemoglobin (HbA1c) in humans: review of recent progress and prospects for use in diabetes care. *Curr Diab Rep.* 2011;11(6):562-569.
22. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: a transethnic genome-wide meta-analysis. *PLoS Med.* 2017;14(9):e1002383.
23. Svensson J, Cerqueira C, Kjaergaard P, et al. Danish registry of childhood and adolescent diabetes. *Clin Epidemiol.* 2016;8:679-683.
24. Fredheim S, Johannesen J, Johansen A, et al. Diabetic ketoacidosis at the onset of type 1 diabetes is associated with future HbA1c levels. *Diabetologia.* 2013;56(5):995-1003.
25. Olsen B, Johannesen J, Fredheim S, Svensson J. Insulin pump treatment; increasing prevalence, and predictors for better metabolic outcome in Danish children and adolescents with type 1 diabetes. *Pediatr Diabetes.* 2015;16(4):256-262.
26. Grabert M, Schweiggert F, Holl RW. A framework for diabetes documentation and quality management in Germany: 10 years of experience with DPV. *Comput Methods Programs Biomed.* 2002;69(2):115-121.
27. Hofer SE, Raile K, Frohlich-Reiterer E, et al. Tracking of metabolic control from childhood to young adulthood in type 1 diabetes. *J Pediatr.* 2014;165(5):956-961 e951-952
28. Konrad K, Datz N, Engelsberger I, et al. Current use of metformin in addition to insulin in pediatric patients with type 1 diabetes mellitus: an analysis based on a large diabetes registry in Germany and Austria. *Pediatr Diabetes.* 2015;16(7):529-537.
29. Rosenbauer J, Dost A, Karges B, et al. Improved metabolic control in children and adolescents with type 1 diabetes: a trend analysis using prospective multicenter data from Germany and Austria. *Diabetes Care.* 2012;35(1):80-86.
30. Scheuing N, Wiegand S, Bachle C, et al. Impact of maternal country of birth on type-1-diabetes therapy and outcome in 27,643 children and adolescents from the DPV registry. *PLoS One.* 2015;10(8):e0135178.
31. Beck RW, Tamborlane WV, Bergenstal RM, Miller KM, DuBose SN, Hall CA. The T1D exchange clinic registry. *J Clin Endocrinol Metab.* 2012;97(12):4383-4389.
32. Blackman SM, Raghinaru D, Adi S, et al. Insulin pump use in young children in the T1D exchange clinic registry is associated with lower hemoglobin A1c levels than injection therapy. *Pediatr Diabetes.* 2014;15(8):564-572.
33. Campbell MS, Schatz DA, Chen V, et al. A contrast between children and adolescents with excellent and poor control: the T1D exchange clinic registry experience. *Pediatr Diabetes.* 2014;15(2):110-117.
34. Cengiz E, Xing D, Wong JC, et al. Severe hypoglycemia and diabetic ketoacidosis among youth with type 1 diabetes in the T1D exchange clinic registry. *Pediatr Diabetes.* 2013;14(6):447-454.
35. Clements MA, Foster NC, Maahs DM, et al. Hemoglobin A1c (HbA1c) changes over time among adolescent and young adult participants in the T1D exchange clinic registry. *Pediatr Diabetes.* 2016;17(5):327-336.
36. Libman IM, Miller KM, DiMeglio LA, et al. Effect of metformin added to insulin on glycemic control among overweight/obese adolescents with type 1 diabetes: a randomized clinical trial. *JAMA.* 2015;314(21):2241-2250.
37. Willi SM, Miller KM, DiMeglio LA, et al. Racial-ethnic disparities in management and outcomes among children with type 1 diabetes. *Pediatrics.* 2015;135(3):424-434.
38. Hanberger L, Samuelsson U, Lindblad B, Ludvigsson J. A1C in children and adolescents with diabetes in relation to certain clinical parameters: the Swedish Childhood Diabetes Registry SWEDIABKIDS. *Diabetes Care.* 2008;31(5):927-929.
39. Beraki A, Magnuson A, Sarnblad S, Aman J, Samuelsson U. Increase in physical activity is associated with lower HbA1c levels in children and adolescents with type 1 diabetes: results from a cross-sectional study based on the Swedish pediatric diabetes quality registry (SWEDIABKIDS). *Diabetes Res Clin Pract.* 2014;105(1):119-125.
40. Hanberger L, Akesson K, Samuelsson U. Glycated haemoglobin variations in paediatric type 1 diabetes: the impact of season, gender and age. *Acta Paediatr.* 2014;103(4):398-403.
41. Nylander C, Toivonen H, Nasic S, Soderstrom U, Tindberg Y, Fernell E. Children and adolescents with type 1 diabetes and high HbA1c—a neurodevelopmental perspective. *Acta Paediatr.* 2013;102(4):410-415.
42. Bachle C, Lange K, Stahl-Pehe A, et al. Associations between HbA1c and depressive symptoms in young adults with early-onset type 1 diabetes. *Psychoneuroendocrinology.* 2015;55:48-58.
43. Bachle C, Stahl-Pehe A, Rosenbauer J. Disordered eating and insulin restriction in youths receiving intensified insulin treatment: results from a nationwide population-based study. *Int J Eat Disord.* 2016;49(2):191-196.
44. Ogilvie D, Fayer D, Petticrew M, et al. The harvest plot: a method for synthesising evidence about the differential effects of interventions. *BMC Med Res Methodol.* 2008;8(1):8.
45. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014;42(Database issue):D1001-D1006.
46. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43(Database issue):D447-D452.
47. Cadario F, Cerutti F, Savastio S, Rabbone I, Tumini S, Bruno G. Increasing burden, younger age at onset and worst metabolic control in migrant than in Italian children with type 1 diabetes: an emerging problem in pediatric clinics. *Acta Diabetol.* 2014;51(2):263-267.
48. Chemtob CM, Hochhauser CJ, Shemesh E, Schmeidler J, Rapoport R. Does poor early metabolic control predict subsequent poor control in young children with type 1 diabetes: an exploratory study. *J Diabetes.* 2011;3(2):153-157.
49. Cherubini V, Pintaudi B, Rossi MC, et al. Severe hypoglycemia and ketoacidosis over one year in Italian pediatric population with type 1 diabetes mellitus: a multicenter retrospective observational study. *Nutr Metab Cardiovasc Dis.* 2014;24(5):538-546.
50. de Beaufort CE, Lange K, Swift PG, et al. Metabolic outcomes in young children with type 1 diabetes differ between treatment centers: the Hvidoere Study in Young Children 2009. *Pediatr Diabetes.* 2013;14(6):422-428.
51. Deladoey J, Henderson M, Geoffroy L. Linear association between household income and metabolic control in children with insulin-dependent diabetes mellitus despite free access to health care. *J Clin Endocrinol Metab.* 2013;98(5):E882-E885.
52. Fendler W, Baranowska AI, Mianowska B, Szadkowska A, Mlynarski W. Three-year comparison of subcutaneous insulin pump treatment with multi-daily injections on HbA1c, its variability and hospital burden of children with type 1 diabetes. *Acta Diabetol.* 2012;49(5):363-370.
53. Hilmi A, Pasternak Y, Friger M, Loewenthal N, Haim A, Hershkovitz E. Ethnic differences in glycemic control and diabetic ketoacidosis rate among children with diabetes mellitus type 1 in the Negev area. *Isr Med Assoc J.* 2013;15(6):267-270.
54. Jescic M, Sajic S, Bogicevic D, Buljugic S, Maglajlic S. The effect of metabolic and hormonal parameters on microalbuminuria in adolescents with type 1 diabetes mellitus. *Srp Arh Celok Lek.* 2013;141(5-6):315-319.
55. Kalyva E, Abdul-Rasoul M, Kehl D, Barkai L, Lukacs A. A cross-cultural study on perceived health-related quality of life in children and adolescents with type 1 diabetes mellitus. *J Diabetes Complications.* 2016;30(3):482-487.
56. King PS, Berg CA, Butner J, et al. Longitudinal trajectories of metabolic control across adolescence: associations with parental involvement, adolescents' psychosocial maturity, and health care utilization. *J Adolesc Health.* 2012;50(5):491-496.



57. Law GU, Walsh J, Queralt V, Nouwen A. Adolescent and parent diabetes distress in type 1 diabetes: the role of self-efficacy, perceived consequences, family responsibility and adolescent-parent discrepancies. *J Psychosom Res.* 2013;74(4):334-339.
58. Marcovecchio ML, Heywood JJ, Dalton RN, Dunger DB. The contribution of glycemic control to impaired growth during puberty in young people with type 1 diabetes and microalbuminuria. *Pediatr Diabetes.* 2014;15(4):303-308.
59. Mauras N, Beck R, Xing D, et al. A randomized clinical trial to assess the efficacy and safety of real-time continuous glucose monitoring in the management of type 1 diabetes in young children aged 4 to <10 years. *Diabetes Care.* 2012;35(2):204-210.
60. Nadeau KJ, Chow K, Alam S, et al. Effects of low dose metformin in adolescents with type 1 diabetes mellitus: a randomized, double-blinded placebo-controlled study. *Pediatr Diabetes.* 2015;16(3):196-203.
61. Petit-Bibal C, Rothenbuhler A, Lucchini P, et al. Decrease in clinical hypoglycemia in young children with type 1 diabetes treated with free-mixed aspart and detemir insulin: an open labeled randomized trial. *Pediatr Diabetes.* 2015;16(3):345-353. Available from: <http://onlinelibrary.wiley.com/doi/10.1111/1469-7610.12437>
62. Rabbone I, Scaramuzza AE, Ignaccolo MG, et al. Carbohydrate counting with an automated bolus calculator helps to improve glycaemic control in children with type 1 diabetes using multiple daily injection therapy: an 18-month observational study. *Diabetes Res Clin Pract.* 2014;103(3):388-394.
63. Redon I, Bertrand J, Martin D, et al. Changes in insulin therapy regimens over 10 yr in children and adolescents with type 1 diabetes attending diabetes camps. *Pediatr Diabetes.* 2014;15(5):329-335.
64. Redondo MJ, Connor CG, Ruedy KJ, et al. Pediatric Diabetes Consortium Type 1 Diabetes New Onset (NeOn) Study: factors associated with HbA1c levels one year after diagnosis. *Pediatr Diabetes.* 2014;15(4):294-302.
65. Sood ED, Pendley JS, Delamater AM, Rohan JM, Pulgaron ER, Drotar D. Mother-father informant discrepancies regarding diabetes management: associations with diabetes-specific family conflict and glycemic control. *Health Psychol.* 2012;31(5):571-579.
66. Stahl-Pehe A, Strassburger K, Castillo K, et al. Quality of life in intensively treated youths with early-onset type 1 diabetes: a population-based survey. *Pediatr Diabetes.* 2014;15(6):436-443.
67. Telo GH, Volkening LK, Butler DA, Laffel LM. Salient characteristics of youth with type 1 diabetes initiating continuous glucose monitoring. *Diabetes Technol Ther.* 2015;17(6):373-378.
68. Valenzuela JM, Smith LB, Stafford JM, et al. Shared decision-making among caregivers and health care providers of youth with type 1 diabetes. *J Clin Psychol Med Settings.* 2014;21(3):234-243.
69. Waller D, Johnston C, Molyneux L, et al. Glycemic control and blood glucose monitoring over time in a sample of young Australians with type 1 diabetes: the role of personality. *Diabetes Care.* 2013;36(10):2968-2973.
70. Abali S, Turan S, Atay Z, Guran T, Haliloglu B, Bereket A. Higher insulin detemir doses are required for the similar glycemic control: comparison of insulin detemir and glargine in children with type 1 diabetes mellitus. *Pediatr Diabetes.* 2015;16(5):361-366.
71. Baucom KJ, Queen TL, Wiebe DJ, et al. Depressive symptoms, daily stress, and adherence in late adolescents with type 1 diabetes. *Health Psychol.* 2015;34(5):522-530.
72. Bizzarri C, Benevento D, Patera IP, et al. Residual beta-cell mass influences growth of prepubertal children with type 1 diabetes. *Horm Res Paediatr.* 2013;80(4):287-292.
73. Brancato D, Fleres M, Aiello V, et al. The effectiveness and durability of an early insulin pump therapy in children and adolescents with type 1 diabetes mellitus. *Diabetes Technol Ther.* 2014;16(11):735-741.
74. Brorsson AL, Viklund G, Orqvist E, Lindholm Olinder A. Does treatment with an insulin pump improve glycaemic control in children and adolescents with type 1 diabetes? A retrospective case-control study. *Pediatr Diabetes.* 2015;16(7):546-553.
75. Butwicka A, Zalepa A, Fendler W, Szadkowska A, Mlynarski W. Maternal depressive symptoms predict acute hospitalization among children with type 1 diabetes. *Pediatr Diabetes.* 2013;14(4):288-294.
76. Coulon SJ, Velasco-Gonzalez C, Scribner R, et al. Racial differences in neighborhood disadvantage, inflammation and metabolic control in black and white pediatric type 1 diabetes patients. *Pediatr Diabetes.* 2017;18(2):120-127.
77. Danne T, Philotheou A, Goldman D, et al. A randomized trial comparing the rate of hypoglycemia—assessed using continuous glucose monitoring—in 125 preschool children with type 1 diabetes treated with insulin glargine or NPH insulin (the PRESCHOOL study). *Pediatr Diabetes.* 2013;14(4):288-294. Available from: <http://onlinelibrary.wiley.com/doi/10.1111/1469-7610.12437>
78. de Bock M, Gunn AJ, Holt JA, et al. Impact of insulin pumps on glycaemic control in a pump-naive paediatric regional population. *J Paediatr Child Health.* 2012;48(3):247-252.
79. Demirel F, Tepe D, Esen I, Buber N, Boztepe H. Individual and familial factors associated with metabolic control in children with type 1 diabetes. *Pediatr Int.* 2013;55(6):710-713.
80. Drew LM, Berg C, King P, et al. Depleted parental psychological resources as mediators of the association of income with adherence and metabolic control. *J Fam Psychol.* 2011;25(5):751-758.
81. Fortenberry KT, Berg CA, King PS, et al. Longitudinal trajectories of illness perceptions among adolescents with type 1 diabetes. *J Pediatr Psychol.* 2014;39(7):687-696.
82. Galler A, Lindau M, Ernert A, Thalemann R, Raile K. Associations between media consumption habits, physical activity, socioeconomic status, and glycemic control in children, adolescents, and young adults with type 1 diabetes. *Diabetes Care.* 2011;34(11):2356-2359.
83. Goksen D, Atik Altinok Y, Ozen S, Demir G, Darcan S. Effects of carbohydrate counting method on metabolic control in children with type 1 diabetes mellitus. *J Clin Res Pediatr Endocrinol.* 2014;6(2):74-78.
84. Helgeson VS, Becker D, Escobar O, Siminerio L. Families with children with diabetes: implications of parent stress for parent and child health. *J Pediatr Psychol.* 2012;37(4):467-478.
85. Hilliard ME, Holmes CS, Chen R, Maher K, Robinson E, Streisand R. Disentangling the roles of parental monitoring and family conflict in adolescents' management of type 1 diabetes. *Health Psychol.* 2013;32(4):388-396.
86. Hilliard ME, Rohan JM, Carle AC, Pendley JS, Delamater A, Drotar D. Fathers' involvement in preadolescents' diabetes adherence and glycemic control. *J Pediatr Psychol.* 2011;36(8):911-922.
87. Hughes AE, Berg CA, Wiebe DJ. Emotional processing and self-control in adolescents with type 1 diabetes. *J Pediatr Psychol.* 2012;37(8):925-934.
88. Jinno K, Urakami T, Horikawa R, et al. Usefulness of insulin detemir in Japanese children with type 1 diabetes. *Pediatr Int.* 2012;54(6):773-779.
89. Katz ML, Mehta S, Nansel T, Quinn H, Lipsky LM, Laffel LM. Associations of nutrient intake with glycemic control in youth with type 1 diabetes: differences by insulin regimen. *Diabetes Technol Ther.* 2014;16(8):512-518.
90. Kim H, Elmi A, Henderson CL, Cogen FR, Kaplowitz PB. Characteristics of children with type 1 diabetes and persistent suboptimal glycemic control. *J Clin Res Pediatr Endocrinol.* 2012;4(2):82-88.
91. Lukacs A, Kiss-Toth E, Varga B, Soos A, Takacs P, Barkai L. Benefits of continuous subcutaneous insulin infusion on quality of life. *Int J Technol Assess Health Care.* 2013;29(1):48-52.
92. Mackey ER, Struempf K, Powell PW, Chen R, Streisand R, Holmes CS. Maternal depressive symptoms and disease care status in youth with type 1 diabetes. *Health Psychol.* 2014;33(8):783-791.

93. Mameli C, Scaramuzza AE, Ho J, Cardona-Hernandez R, Suarez-Ortega L, Zuccotti GV. A 7-year follow-up retrospective, international, multicenter study of insulin pump therapy in children and adolescents with type 1 diabetes. *Acta Diabetol.* 2014;51(2):205-210.
94. Mianowska B, Kaminska A, Fendler W, Szadkowska A, Mlynarski W. Bilirubin is an independent factor inversely associated with glycosylated hemoglobin level in pediatric patients with type 1 diabetes. *Pediatr Diabetes.* 2014;15(5):389-393.
95. Nansel TR, Lipsky LM, Iannotti RJ. Cross-sectional and longitudinal relationships of body mass index with glycemic control in children and adolescents with type 1 diabetes mellitus. *Diabetes Res Clin Pract.* 2013;100(1):126-132.
96. Price KJ, Knowles JA, Fox M, et al. Effectiveness of the Kids in Control of Food (KICK-OFF) structured education course for 11-16 year olds with type 1 diabetes. *Diabet Med.* 2016;33(2):192-203.
97. Pulgaron ER, Sanders LM, Patino-Fernandez AM, et al. Glycemic control in young children with diabetes: the role of parental health literacy. *Patient Educ Couns.* 2014;94(1):67-70.
98. Rohan JM, Rausch JR, Pendley JS, et al. Identification and prediction of group-based glycemic control trajectories during the transition to adolescence. *Health Psychol.* 2014;33(10):1143-1152.
99. Sands AL, Higgins LA, Mehta SN, Nansel TR, Lipsky LM, Laffel LM. Associations of youth and parent weight status with reported versus predicted daily energy intake and hemoglobin A1c in youth with type 1 diabetes mellitus. *J Diabetes Sci Technol.* 2013;7(1):263-270.
100. Shalitin S, Phillip M. Which factors predict glycemic control in children diagnosed with type 1 diabetes before 6.5 years of age? *Acta Diabetol.* 2012;49(5):355-362.
101. Shorer M, David R, Schoenberg-Taz M, Levavi-Lavi I, Phillip M, Meyerovitch J. Role of parenting style in achieving metabolic control in adolescents with type 1 diabetes. *Diabetes Care.* 2011;34(8):1735-1737.
102. Wang JT, Wiebe DJ, White PC. Developmental trajectories of metabolic control among White, Black, and Hispanic youth with type 1 diabetes. *J Pediatr.* 2011;159(4):571-576.
103. Wiebe DJ, Gelfand D, Butler JM, et al. Longitudinal associations of maternal depressive symptoms, maternal involvement, and diabetes management across adolescence. *J Pediatr Psychol.* 2011;36(7):837-846.
104. Wisting L, Bang L, Natvig H, et al. Metabolic control and illness perceptions in adolescents with type 1 diabetes. *J Diabet Res.* 2016;2016. 3486094
105. Wong JC, Dolan LM, Yang TT, Hood KK. Insulin pump use and glycemic control in adolescents with type 1 diabetes: predictors of change in method of insulin delivery across two years. *Pediatr Diabetes.* 2015;16(8):592-599.
106. Wu YP, Hilliard ME, Rausch J, Dolan LM, Hood KK. Family involvement with the diabetes regimen in young people: the role of adolescent depressive symptoms. *Diabetic Medicine: A Journal of the British Diabetic Association.* 2013;30(5):596-602.
107. Zdunczyk B, Sendela J, Zypowska A. High prevalence of depressive symptoms in well-controlled adolescents with type 1 diabetes treated with continuous subcutaneous insulin infusion. *Diabetes Metab Res Rev.* 2014;30(4):333-338.
108. Miller KM, Foster NC, Beck RW, et al. Current state of type 1 diabetes treatment in the U.S.: updated data from the T1D exchange clinic registry. *Diabetes Care.* 2015;38(6):971-978.
109. Lampert T, Sygusch R, Schlack R. Use of electronic media in adolescence. Results of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2007;50(5-6):643-652.
110. Freidlin B, Zheng G, Li Z, Gastwirth JL. Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum Hered.* 2002;53(3):146-152.
111. Kaiserman K, Jung H, Benabbad I, Karges B, Polak M, Rosilio M. 20 years of insulin lispro in pediatric type 1 diabetes: a review of available evidence. *Pediatr Diabetes.* 2017;18(2):81-94.
112. La Greca AM, Follansbee D, Skyler JS. Developmental and behavioral aspects of diabetes management in youngsters. *Child Health Care.* 1990;19(3):132-139.
113. Iannotti RJ, Nansel TR, Schneider S, et al. Assessing regimen adherence of adolescents with type 1 diabetes. *Diabetes Care.* 2006;29(10):2263-2267.
114. Harris MA, Wysocki T, Sadler M, et al. Validation of a structured interview for the assessment of diabetes self-management. *Diabetes Care.* 2000;23(9):1301-1304.
115. Holmes CS, Chen R, Streisand R, et al. Predictors of youth diabetes care behaviors and metabolic control: a structural equation modeling approach. *J Pediatr Psychol.* 2006;31(8):770-784.
116. Young MT, Lord JH, Patel NJ, Gruhn MA, Jaser SS. Good cop, bad cop: quality of parental involvement in type 1 diabetes management in youth. *Curr Diab Rep.* 2014;14(11):546.
117. Must A, Anderson SE. Body mass index in children and adolescents: considerations for population-based applications. *Int J Obes (Lond).* 2006;30(4):590-594.
118. Buchberger B, Huppertz H, Krabbe L, Lux B, Mattivi JT, Siafarikas A. Symptoms of depression and anxiety in youth with type 1 diabetes: a systematic review and meta-analysis. *Psychoneuroendocrinology.* 2016;70:70-84.
119. Rechenberg K, Whittemore R, Grey M. Anxiety in youth with type 1 diabetes. *J Pediatr Nurs.* 2017;32:64-71.
120. Petrone A, Spoletini M, Zampetti S, et al. The PTPN22 1858T gene variant in type 1 diabetes is associated with reduced residual beta-cell function and worse metabolic control. *Diabetes Care.* 2008;31(6):1214-1218.
121. Nielsen LB, Porksen S, Andersen ML, et al. The PTPN22 C1858T gene variant is associated with proinsulin in new-onset type 1 diabetes. *BMC Med Genet.* 2011;12(1):41.
122. Mosaad YM, Elsharkawy AA, El-Deek BS. Association of CTLA-4 (+49A/G) gene polymorphism with type 1 diabetes mellitus in Egyptian children. *Immunol Invest.* 2012;41(1):28-37.
123. Almasi S, Aliparasti MR, Aliasgarzadeh A, et al. Investigation of CTLA-4-318C/T gene polymorphism in cases with type 1 diabetes of Azerbaijan, Northwest Iran. *Immunol Lett.* 2015;166(2):134-139.
124. Assmann TS, Brondani Lde A, Bauer AC, Canani LH, Crispim D. Polymorphisms in the TLR3 gene are associated with risk for type 1 diabetes mellitus. *Eur J Endocrinol.* 2014;170(4):519-527.
125. Spoletini M, Petrone A, Zampetti S, et al. Low-risk HLA genotype in type 1 diabetes is associated with less destruction of pancreatic B-cells 12 months after diagnosis. *Diabet Med.* 2007;24(12):1487-1490.
126. Korolija M, Hadzija M, Medvidovic EP, et al. Genetic evaluation of the TNF-alpha -238G>A and -308G>A promoter polymorphisms in Croatian patients with type 1 diabetes. *Hum Immunol.* 2010;71(12):1228-1232.
127. Ururahy MA, de Souza KS, Oliveira YM, et al. Association of polymorphisms in IL6 gene promoter region with type 1 diabetes and increased albumin-to-creatinine ratio. *Diabetes Metab Res Rev.* 2015;31(5):500-506.
128. Kaur S, Mirza AH, Brorsson CA, et al. The genetic and regulatory architecture of ERBB3-type 1 diabetes susceptibility locus. *Mol Cell Endocrinol.* 2016;419:83-91.
129. Bouhanick B, Gallois Y, Hadjadj S, Boux de Casson F, Limal JM, Marre M. Relationship between glomerular hyperfiltration and ACE insertion/deletion polymorphism in type 1 diabetic children and adolescents. *Diabetes Care.* 1999;22(4):618-622.
130. Jacobson AM, Paterson AD, Ryan CM, et al. The associations of apolipoprotein E and angiotensin-converting enzyme polymorphisms and cognitive function in type 1 diabetes based on an 18-year follow-up of the DCCT cohort. *Diabet Med.* 2010;27(1):15-22.



131. Johannesen J, Svensson J, Bergholdt R, et al. Hypoglycemia, S-ACE and ACE genotypes in a Danish nationwide population of children and adolescents with type 1 diabetes. *Pediatr Diabetes*. 2011;12(2):100-106.
132. Brorsson CA, Nielsen LB, Andersen ML, et al. Genetic risk score modelling for disease progression in new-onset type 1 diabetes patients: increased genetic load of islet-expressed and cytokine-regulated candidate genes predicts poorer glycemic control. *J Diabet Res*. 2016;2016:1-8.
133. Mortensen HB, Hougaard P, Swift P, et al. New definition for the partial remission period in children and adolescents with type 1 diabetes. *Diabetes Care*. 2009;32(8):1384-1390.
134. Paterson AD, Waggott D, Borchard AP, et al. A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both A1C and glucose. *Diabetes*. 2010;59(2):539-549.
135. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes*. 2010;59(12):3229-3239.
136. Clee SM, Yandell BS, Schueler KM, et al. Positional cloning of Sorcs1, a type 2 diabetes quantitative trait locus. *Nat Genet*. 2006;38(6):688-693.
137. Granhall C, Park H-B, Fakhrai-Rad H, Luthman H. High-resolution quantitative trait locus analysis reveals multiple diabetes susceptibility loci mapped to intervals <800 kb in the species-conserved Niddm1i of the GK rat. *Genetics*. 2006;174(3):1565-1572.
138. Kebede MA, Oler AT, Gregg T, et al. SORCS1 is necessary for normal insulin secretory granule biogenesis in metabolically stressed beta cells. *J Clin Invest*. 2014;124(10):4240-4256.
139. Lachin JM, McGee P, Palmer JP. Impact of C-peptide preservation on metabolic and clinical outcomes in the Diabetes Control and Complications Trial. *Diabetes*. 2014;63(2):739-748.
140. Cho YH, Craig ME, Donaghue KC. Puberty as an accelerator for diabetes complications. *Pediatr Diabetes*. 2014;15(1):18-26.
141. Schweyer AC, Jellimann S. Adolescence and diabetes, psychological and physiological aspects. *Rev Infirm*. 2015;(216):37-38.
142. Fröhlich-Reiterer EE, Rosenbauer J, Bechtold-Dalla Pozza S, Hofer SE, Schober E, Holl RW. Predictors of increasing BMI during the course of diabetes in children and adolescents with type 1 diabetes: data from the German/Austrian DPV multicentre survey. *Arch Dis Child*. 2014;99(8):738-743.
143. Akesson K, Hanberger L, Samuelsson U. The influence of age, gender, insulin dose, BMI, and blood pressure on metabolic control in young patients with type 1 diabetes. *Pediatr Diabetes*. 2015;16(8):581-586.
144. Hertel JK, Johansson S, Raeder H, et al. Evaluation of four novel genetic variants affecting hemoglobin A1c levels in a population-based type 2 diabetes cohort (the HUNT2 study). *BMC Med Genet*. 2011;12(1):20.
145. An P, Mijlkovic I, Thyagarajan B, et al. Genome-wide association study identifies common loci influencing circulating glycated hemoglobin (HbA1c) levels in non-diabetic subjects: the Long Life Family Study (LLFS). *Metabolism*. 2014;63(4):461-468.
146. Bonnefond A, Vaxillaire M, Labruno Y, et al. Genetic variant in HK1 is associated with a proanemic state and A1C but not other glycemic control-related traits. *Diabetes*. 2009;58(11):2687-2697.
147. Chen P, Ong RT, Tay WT, et al. A study assessing the association of glycated hemoglobin A1C (HbA1C) associated variants with HbA1C, chronic kidney disease and diabetic retinopathy in populations of Asian ancestry. *PLoS One*. 2013;8(11):e79767.
148. Chen P, Takeuchi F, Lee JY, et al. Multiple nonglycemic genomic loci are newly associated with blood level of glycated hemoglobin in East Asians. *Diabetes*. 2014;63(7):2551-2562.
149. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42(2):105-116.
150. Franklin CS, Aulchenko YS, Huffman JE, et al. The TCF7L2 diabetes risk variant is associated with HbA(1)(C) levels: a genome-wide association meta-analysis. *Ann Hum Genet*. 2010;74(6):471-478.
151. Pare G, Chasman DI, Parker AN, et al. Novel association of HK1 with glycated hemoglobin in a non-diabetic population: a genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. *PLoS Genet*. 2008;4(12):e1000312.
152. Ryu J, Lee C. Association of glycosylated hemoglobin with the gene encoding CDKAL1 in the Korean Association Resource (KARE) study. *Hum Mutat*. 2012;33(4):655-659.
153. Appel EVR, Moltke I, Jorgensen ME, et al. Genetic determinants of glycated hemoglobin levels in the Greenlandic Inuit population. *Eur J Hum Genet*. 2018;26(6):868-875.
154. Maahs DM, Hermann JM, DuBose SN, et al. Contrasting the clinical care and outcomes of 2,622 children with type 1 diabetes less than 6 years of age in the United States T1D exchange and German/Austrian DPV registries. *Diabetologia*. 2014;57(8):1578-1585.
155. Maahs DM, Hermann JM, Holman N, et al. Rates of diabetic ketoacidosis: international comparison with 49,859 pediatric patients with type 1 diabetes from England, Wales, the U.S., Austria, and Germany. *Diabetes Care*. 2015;38(10):1876-1882.
156. Peterson A, Hanberger L, Akesson K, Bojestig M, Andersson Gare B, Samuelsson U. Improved results in paediatric diabetes care using a quality registry in an improvement collaborative: a case study in Sweden. *PLoS One*. 2014;9(5):e97875.
157. Moosavi M, Seguin J, Polychronakos C. Effect of autoimmunity risk loci on the honeymoon phase in type 1 diabetes. *Pediatr Diabetes*. 2017;18(6):459-462.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Gloaguen E, Bendelac N, Nicolino M, Julier C, Mathieu F. A systematic review of non-genetic predictors and genetic factors of glycated haemoglobin in type 1 diabetes one year after diagnosis. *Diabetes Metab Res Rev*. 2018;34:e3051. <https://doi.org/10.1002/dmrr.3051>

# Chapitre 4 | Recherche de facteurs influençant le taux d'HbA1c dans une cohorte d'enfants français atteints de DT1 – Etude préliminaire

---

## 4.1 Introduction

La revue de la littérature, présentée dans le Chapitre 3, montre que l'équilibre glycémique chez l'enfant souffrant de DT1 est un trait complexe influencé par de nombreux facteurs, qui sont démographiques, cliniques ou génétiques. Ces facteurs peuvent être eux-mêmes corrélés entre eux, rendant leur analyse complexe (Gloaguen *et al.*, 2018). Cette revue a aussi permis de mettre en évidence le peu d'études à ce moment-là s'intéressant aux facteurs génétiques du taux d'HbA1c chez le patient souffrant de DT1.

Depuis, un certain nombre de nouvelles études ont été publiées sur les mêmes facteurs non-génétiques, en essayant notamment de mieux caractériser le lien entre certains facteurs. Par exemple, un nombre assez important d'études ont essayé de mieux comprendre le lien entre l'appartenance à une minorité ethnique et le statut socioéconomique du patient et l'impact de ces deux facteurs sur le taux d'HbA1c (Agarwal *et al.*, 2019; Bergenstal *et al.*, 2017; Mönkemöller *et al.*, 2019; Thomas *et al.*, 2018; Zuckerman-Levin *et al.*, 2018).

D'un point de vue génétique, depuis la publication de la revue, cinq études ont été publiées ; quatre études utilisent une stratégie de gènes-candidats (Abu El-Ella *et al.*, 2018; Rokamp *et al.*, 2018; Slominski *et al.*, 2019; Souza *et al.*, 2018) et une stratégie génome-entier (Syreeni *et al.*, 2019). Parmi les études en gènes-candidats, deux ont montré un effet significatif du marqueur testé sur le taux d'HbA1c : les patients ayant le génotype CT et TT du marqueur rs3736228 du gène LRP5 ont en moyenne un taux d'HbA1c plus élevé que les porteurs du génotype CC (Abu

El-Ella *et al.*, 2018) ; les porteurs de la délétion  $\Delta 32$  (rs333) à l'état homozygote ou hétérozygote du gène CCR5 ont en moyenne un taux d'HbA1c plus élevé (Slominski *et al.*, 2019).

Un GWAS a été effectué sur une cohorte de patients souffrant de DT1 (Syreeni *et al.*, 2019). Cette étude est basée sur les données de l'étude FinnDiane (*Finnish Diabetic Nephropathy*), étude multicentrique nationale finnoise ayant débutée en 1997, qui avait pour but d'identifier les facteurs de risques et les mécanismes liés au développement des complications diabétiques. Ce GWAS a été réalisé en prenant comme covariable le genre, l'âge, la durée du diabète, le nombre de mesure du taux d'HbA1c disponibles pour le patient et les critères de qualité du génotypage (lot de génotypage et les 10 premières composantes principales), sous un modèle additif. Les auteurs ont reporté une région significativement associée au taux d'HbA1c sur le chromosome 13 près du gène RXFP2. Dans un second temps, l'association entre cette région et la dose d'insuline requise a été testée, et aucune association significative n'a été montrée. Cette région a été répliquée dans la cohorte *Renin–Angiotensin System Study* (RASS) et dans une méta-analyse, mais à cause de la fréquence faible de ces variants, l'association n'a pas pu être testée dans toutes les cohortes de patients DT1 disponibles.

Au cours de ma thèse, j'ai participé au projet GENDIR, pour « Étude des facteurs GENétiques impliqués dans le contrôle de la Dose d'Insuline journalière Requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type 1 », qui vise à étudier le rôle des facteurs impliqués dans le contrôle de l'équilibre glycémique dans le cadre d'une étude rétrospective française. Les objectifs du projet GENDIR sont de rechercher des facteurs de risque (génétiques ou non) impliqués dans le contrôle de l'équilibre glycémique en ajustant sur l'effet du traitement et des autres covariables et d'étudier l'évolution de l'équilibre glycémique à 5 et 10 ans afin d'identifier les facteurs de risque responsables des différents profils d'évolution de la maladie.

## 4.2 Matériel et méthodes

### 4.2.1 Recrutement des patients et recueil des données

Ce projet est réalisé en collaboration avec le Dr N. Bendelac et le Pr M. Nicolino, chef du service d'endocrinologie-diabétologie pédiatrique de l'Hôpital Femme-Mère-Enfant de Lyon. Cette recherche, sous promotion Inserm, a été classée comme recherche non-interventionnelle et a obtenu les différentes autorisations nécessaires pour pouvoir commencer le 16 février 2015.

Les critères d'inclusion des patients pour le projet GENDIR sont :

- (1) Souffrir de DT1 ;
- (2) Avoir une lignée ou un échantillon d'ADN disponible dans la collection biologique hébergée par le Centre de Biotechnologie Cellulaire et Biothèque (CBC-Biotec) des Hospices Civils de Lyon ;
- (3) Avoir au moins une valeur de la dose d'insuline requise et un dosage du taux d'HbA1c après 2 ans de maladie ;
- (4) Avoir signé un consentement éclairé ou l'avoir fait signer par les parents du patient pour un sujet mineur.

L'ensemble des patients incluables a été recontacté entre mars 2015 et août 2017, afin d'obtenir leur consentement écrit. Trois types de consentements ont été utilisés en fonction de l'âge du patient au moment du recrutement : un consentement pour les patients âgés entre 6 et 10 ans au moment de l'étude, un pour les patients âgés entre 11 et 17 ans et un dernier pour les patients adultes (Annexe 3). Aucun patient âgé de moins de 6 ans n'a été contacté puisque les échantillons d'ADN disponibles ont été prélevés entre 2002 et 2009.

Les données ont été anonymisées. Les patients ont été identifiés par un code GENDIR correspondant à L005XXX avec XXX un nombre à trois chiffres correspondant à l'ordre d'inclusion.

L'e-CRF (Cahier de recueil de données électronique, CRF pour l'anglais *Case Report Form*) a été créé à partir du CRF présenté en Annexe 4 et a été développé en utilisant l'application REDCap (pour l'anglais *Research Electronic Data Capture* ; <https://www.project-redcap.org/>) (Harris et al., 2009) dans le cadre d'une collaboration avec l'infrastructure nationale BIOBANQUE

(INSERM US013). L'e-CRF ainsi que la base de données sont hébergés sur le serveur du département système d'information (DSI) de l'Inserm. Le développement de l'e-CRF a été réalisé par Thomas Diot, étudiant en médecine dans le cadre de son stage de M1 Biologie-Santé de l'Université Paris-Est (UPEC) à Paris, sous la direction du Dr. F. Mathieu.

Le recueil des données a été effectué selon deux sources :

- Le système de traitement des données du patient, dans le cadre d'un logiciel développé par les Hospices civils de Lyon, totalement sécurisé et respectant le secret médical et accessible à partir des ordinateurs du service d'endocrinologie pédiatrique ;
- Le dossier médical papier du patient consultable sur demande dans la salle de lecture des archives des Hospices Civils de Lyon.

Durant cette période, 668 courriers ont été envoyés aux patients incluables pour les informer de la recherche et leur proposer d'y participer. 407 patients ont signé un formulaire de consentement entre avril 2015 et août 2017 (Figure 16). Pour 73 patients, le courrier nous a été retourné au motif que le destinataire n'habitait plus à l'adresse indiquée (NPAI). Nous avons eu l'autorisation du comité de protection des personnes d'exploiter leurs données cliniques mais pas leurs données génétiques.

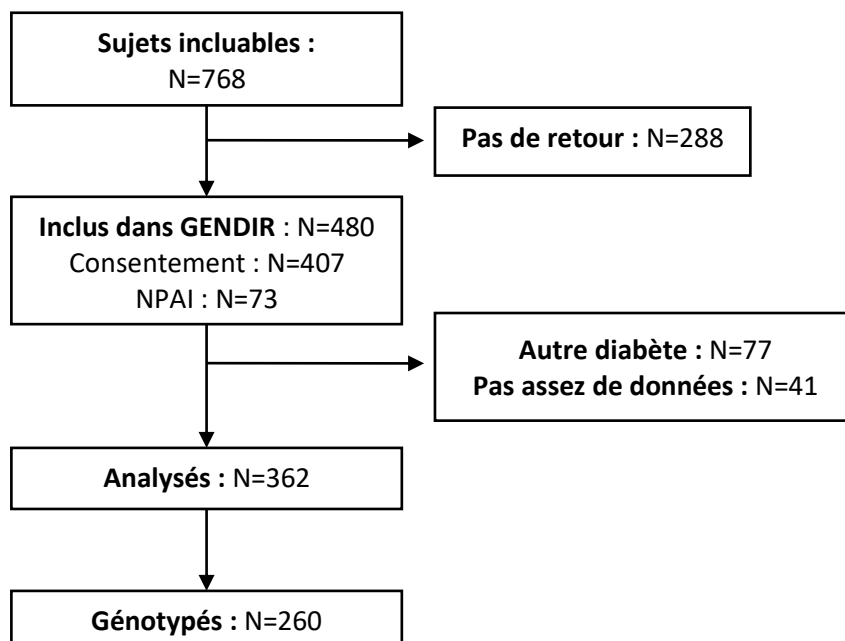


Figure 16 : Flow-chart de l'étude GENDIR

Parmi les 480 patients inclus, 77 patients ont été exclus après lecture de leur dossier médical, car ils avaient été considérés à tort comme souffrant de DT1 (sujet souffrant de diabète apparentés de patient déjà inclus dans l'analyse, sujet atteint de drépanocytose, d'une autre forme de diabète (néonatal, MODY, diabète atypique) ou d'un syndrome particulier (Wolfram), ...) et 41 dossiers n'avaient pas suffisamment de données pour être exploitables.

362 patients souffrant de DT1 ont donc été inclus et seront étudiés dans l'étude (Figure 16). Les statistiques descriptives de l'étude préliminaire de cette cohorte ont été réalisées sur un échantillon de 315 patients souffrant de DT1, pour lesquels la majorité des données avaient été collectées.

#### 4.2.2 Transformation des données

Les patients atteints de DT1 sont en moyenne vus par leur endocrinologue tous les 3 mois dans le cadre du suivi de leur diabète. Les données à 2, 5 et 10 ans (taux d'HbA1c, DIR, IMC) correspondent à la moyenne de l'ensemble des valeurs disponibles dans les dossiers cliniques après 2, 5 et 10 années de diabète plus ou moins 6 mois.

Afin de pouvoir analyser au mieux les données, nous avons décidé de transformer certaines d'entre elles.

##### 4.2.2.1 Les données glycémiques et cliniques

###### IMC standardisé

Chez les enfants, l'utilisation de l'IMC varie beaucoup en fonction de l'âge et du sexe de l'enfant. Il est donc conseillé d'utiliser une valeur d'IMC standardisé (ou IMC-SDS). Pour cela, nous avons utilisé le package *childsds* dans R (Vogel, 2019), en utilisant comme valeurs de référence des données disponibles belges (Roelants, Hauspie, & Hoppenbrouwers, 2009).

###### Insuline Dose Adjusted HbA1c (IDAA1C)

Les données de peptide-C n'étant renseignées que dans 5% des dossiers, nous avons calculé l'HbA1c ajustée pour la dose d'Insuline ou *Insulin Dose Adjusted HbA1c* (IDAA1C), en utilisant la formule développée par Mortensen (Mortensen *et al.*, 2009) et validée (Max Andersen *et al.*, 2014) suivante :  $IDAA1C = HbA1c (\%) + 4 \times DIR (U/kg/Jour)$ , avec la DIR pour la dose



journalière d'insuline requise. Cette mesure reflète la réserve de cellules  $\beta$  des ilots de Langerhans du pancréas.

### Puberté

Nous avons collecté les statuts pubertaires en fonction de la classification de Tanner des patients aux différentes visites à 2, 5 et 10 ans dans les dossiers médicaux. Cependant, cette donnée était rarement renseignée. Nous avons défini avec les cliniciens 3 stades en fonction de l'âge et du genre des patients :

- Enfant :
  - o Pour les filles âgées de moins de 9 ans ;
  - o Pour les garçons âgés de moins de 11 ans ;
- Adolescent :
  - o Pour les filles âgées entre 9 ans et 17 exclus ;
  - o Pour les garçons âgés entre 11 ans et 17 exclus ;
- Adulte : Pour tous les patients de 17 ans et plus.

### Type de traitement

Nous avons considéré deux types de traitement : les patients sous pompe et les patients sous injections, regroupant les patients recevant un traitement « basal-bolus » composé de deux types d'insuline et les patients recevant des mixtes d'insuline.

#### *4.2.2.1 Les données socio-économiques*

##### Catégories socio-professionnel (CSP)

Quand cela était disponible dans les dossiers papiers, la profession des parents au moment du diagnostic a été collectée, puis dans un second temps transformée en catégorie socio-professionnelle (CSP) à l'aide de la nomenclature des professions et catégories socioprofessionnelles des emplois salariés des employeurs privés et publics disponible sur le site de l'Insee (Institut national de la statistique et des études économiques ; [www.insee.fr](http://www.insee.fr)). Huit catégories socioprofessionnelles sont définies par l'Insee :

- Agriculteurs, exploitants : 1

- Artisans, commerçants, chefs d'entreprises : 2
- Cadres et professions intellectuelles supérieures : 3
- Professions intermédiaires : 4
- Employés : 5
- Ouvriers : 6
- Retraités : 7
- Autres personnes sans activité professionnelle : 8.

En raison de la répartition des effectifs, nous avons recodé ces huit catégories en trois :

- CSP+ : correspondant aux catégories 2, 3 et 4 ;
- CSP- : correspondant aux catégories 5 et 6 ;
- Sans emploi : correspondant aux catégories 7 et 8.

Il n'y avait dans notre échantillon aucun patient dont l'un des deux parents était agriculteur ou exploitants ; nous n'avons donc pas ajouté la catégorie 1 ni dans CSP+ ni dans CSP-.

#### Lieu de naissance des parents

Nous avons collecté le lieu de naissance, de chacun des deux parents, disponible dans les dossiers. Les pères sont nés dans 22 pays différents (Algérie, Angola, Arménie, Bénin, Comores, Cambodge, Espagne, France, Grèce, Irlande, Italie, Maroc, Maurice, Portugal, Russie, Sénégal, Soudan, Syrie, Tchad, Tunisie, Turkménistan, Turquie), et les mères dans 14 pays (Algérie, Cambodge, Congo, Comores, Espagne, France, Italie, Maroc, Pologne, Russie, Slovénie, Syrie, Tunisie, Turquie). Nous avons décidé de regrouper les trois pays du Maghreb (Algérie, Maroc, Tunisie) dans une seule catégorie, noté « Maghreb » ; et regrouper tous les autres pays que la France dans la catégorie « Autre » au vu du plus petit effectif de parents provenant de ces pays.

À partir de l'information de ces lieux de naissance, nous avons créé la variable correspondant au fait d'avoir un, deux ou aucun parent(s) né(s) en France.

#### *4.2.2.2 Les données longitudinales*

Avec les cliniciens, nous avons défini deux valeurs seuils qui permettent de définir les patients bien équilibrés, qui ont un taux d'HbA1c inférieur ou égale à 7.5%, et les patients mal équilibrés

avec un taux d'HbA1c supérieur à 9%. Les patients ayant un taux d'HbA1c entre 7.5 et 9% sont considérés comme intermédiaires.

Pour les patients qui possédaient des données à 2, 5 et 10 ans de diabète, nous avons pu définir différents profils d'évolutions, à partir de ces 3 catégories à chaque temps :

- Des patients stables qui vont rester dans une même catégorie aux 3 temps de suivi ;
- Des patients qui ont leur taux d'HbA1c qui augmente entre 2, 5 et 10 ans ;
- Des patients qui ont leur taux d'HbA1c qui diminue entre 2, 5 et 10 ans ;
- Des patients qui vont voir leur taux d'HbA1c fluctuer en diminuant entre 2 et 5 ans puis en ré-augmentant entre 5 et 10 ans ;
- Et inversement en augmentant puis en diminuant.

### 4.2.3 Matériels biologiques et génotypage

Les ADNs issus de sang ou de lignées ont été préparés par le CBC-Biotec des Hospices Civils de Lyon afin d'être génotypés au Centre National de recherche en Génomique Humaine à l'aide de la puce *Illumina's Infinium Global Screening Array-24 v1.0*.

Cette puce est composée de plus de 640 000 marqueurs, qui sont soit utiles pour avoir une couverture optimum pour des GWAS, pour la recherche clinique, c'est-à-dire connus comme associés à une maladie dans différentes bases de données (ExAC (Lek *et al.*, 2016), ClinVar (Landrum *et al.*, 2014), la base de données *Pharmacogenomics Knowledgebase* (PharmGKB) (Whirl-Carrillo *et al.*, 2012), ou la base de données *National Human Genome Research Institute* (NHGRI, <https://www.genome.gov/>)), soit utiles pour le contrôle qualité.

Pour des raisons budgétaires, nous avons choisis de génotyper notre cohorte à l'aide de cette puce performante et peu chère couvrant l'ensemble du génome. Cependant, notre stratégie initiale au vu du nombre de patients génotypés était une stratégie gène-candidats.

Les gènes sélectionnés ont été classés en quatre catégories, de la même façon que dans notre revue de la littérature (Gloaguen *et al.*, 2018) :

- « Immunité » pour les gènes impliqués dans les processus d'immunité et d'inflammation ;
- « Maladie Auto-Immune » pour les gènes impliqués dans l'auto-immunité ;

- « Métabolisme » pour les gènes impliqués dans les voies métaboliques ou les phénotypes associés, concernant les lipides ou le glucose ;
- « Érythrocyte » pour les gènes impliqués dans le cycle de vie des globules rouges.

Les gènes impliqués dans aucune de ces catégories ont été considérés comme appartenant à la catégorie « Autre ». Pour être considéré dans l'une de ces catégories, il faut être associé à au moins deux traits au seuil de  $5E-5$  dans les données de *GWAS Catalog* (MacArthur *et al.*, 2017) et/ou impliqué dans deux voies métaboliques de *Gene Ontology*, disponible dans String (Szklarczyk *et al.*, 2015).

#### 4.2.4 Analyses statistiques

Nous avons décidé d'exclure les patients pour lesquels nous avons la connaissance de consanguinité dans leur famille, notamment dû à des mariages entre cousins germains, à partir de leur dossier médical. En effet, les patients étant issus de mariage consanguins possèdent un taux d'hétérozygotie plus faible que celle attendu, pouvant entraîner des faux-positifs lors de l'analyse d'association génétique (Marees *et al.*, 2018). De plus, dans les populations consanguines, notamment provenant du Maghreb, une forte proportion de DT1 sont en réalité des formes monogéniques qui sont mal diagnostiquées, et qui peuvent avoir des profils glycémiques différents des patients ayant un DT1 (Slingerland, 2006; Zalloua *et al.*, 2008).

##### 4.2.4.1 Analyses des données cliniques à deux ans de diabète

Pour chaque trait, la médiane et les valeurs minimales et maximales sont données pour toutes les données quantitatives et le pourcentage dans chaque catégorie est donné pour les données catégorielles.

Suite à notre revue de la littérature, nous avons montré l'importance d'étudier le taux d'HbA1c en prenant en compte la DIR. Nous avons donc comparé deux manières de prendre en compte le traitement insulínémique pour l'étude du taux d'HbA1c : soit en utilisant l'IDAA1C proposé par Mortensen *et al* en 2009 pour étudier la rémission partielle, qui prend en compte le taux d'HbA1c et la DIR ; soit en étudiant directement le taux d'HbA1c en ajustant sur la DIR. Au vu de la différence de taille d'échantillon entre la cohorte entière et les patients qui ont été génotypés, nous avons décidé de comparer deux façons d'ajuster sur la DIR, soit en ajustant

pour la DIR directement durant l'analyse (HbA1c ajusté) soit en préajustant sur la DIR en utilisant le coefficient de 0.9 obtenu lors de l'analyse univariée sur l'ensemble de l'échantillon. Le taux d'HbA1c préajusté correspond donc pour un patient  $i$  à :

$$HbA1c\_Préajusté_i = HbA1c_i - 0.9 \times DIR_i$$

Cependant le fait de fixer l'effet de la DIR sur le taux d'HbA1c a l'inconvénient de ne pas permettre de tester les interactions qui peuvent exister entre la DIR et plusieurs autres facteurs influençant le taux d'HbA1c.

Pour les analyses de l'IDAA1C, et du taux d'HbA1c préajusté, pour les données qualitatives, les comparaisons du taux d'HbA1c entre les différentes catégories ont été effectuées à l'aide de tests paramétriques, un test de l'écart réduit pour les variables possédant deux catégories et une analyse de la variance (ANOVA) pour celles possédant plus de deux modalités. Pour les données quantitatives, les analyses ont été effectuées à l'aide de régression linéaire. Dans le cas de l'analyse du taux d'HbA1c ajusté, toutes les analyses ont été faites à l'aide d'une régression linéaire incluant la DIR comme covariable.

Pour chaque méthode d'analyse du taux d'HbA1c, nous avons sélectionné les patients dont les données étaient disponibles pour toutes les variables significativement associées au taux d'HbA1c en univariée au seuil  $p < 0.05$  et ayant moins de 25% de données manquantes sur l'ensemble des patients. Nous avons ensuite effectué une régression linéaire multivariée, puis ont été sélectionnées les variables d'intérêt à l'aide d'une régression linéaire pas à pas descendante, avec comme critère de sélection un seuil de signification inférieur à 5% à chaque étape. Les facteurs significatifs au seuil de 5% ont ensuite été pris en compte dans l'analyse génétique.

Les analyses ont été effectuées à l'aide du logiciel R (<http://www.R-project.org>) (Team., 2017) et les figures ont été réalisées à l'aide du package R *ggplot2* (Wickham, 2016).

#### **4.2.4.2 Analyses des données génétiques**

Nous avons utilisé une stratégie gène-candidat, en étudiant les gènes déjà identifiés comme étant associés au taux d'HbA1c. Lorsque le variant connu comme étant associé au taux d'HbA1c n'était pas génotypé dans la puce, nous en avons recherché un en déséquilibre de liaison

( $r^2 > 0.6$ ) qui soit génotypé, à l'aide de LDlink (<https://ldlink.nci.nih.gov/>) (Machiela & Chanock, 2015).

Une étude de puissance a été réalisée sous Quanto 1.2.4 pour un trait quantitatif pour un échantillon de 211 patients non apparentés (pour lesquels nous avons les données génétiques et clinique) et pour une MAF de 5% (<http://biostats.usc.edu/Quanto.html>) (Gauderman, 2002).

Le contrôle qualité a été effectué en suivant les préconisations pour un GWAS publiées en 2018 (Marees *et al.*, 2018). Suite au contrôle qualité, nous avons exclus :

- Les patients ayant un taux de génotypage inférieur à 80% ;
- Les marqueurs ayant :
  - o Un taux de génotypage inférieur à 80% ;
  - o Une MAF inférieure à 5% ;
  - o Une déviation par rapport à l'équilibre d'HW ( $p < 1E-6$ ) ;
  - o Un taux d'hétérozygotie trop important (compris entre  $\pm 3$  écarts types).

Nous avons effectué une analyse en composante principale des données génétiques afin de corriger pour la stratification de population (Marees *et al.*, 2018).

L'analyse par le GRM a été effectuée en utilisant comme covariables les 2 premières composantes principales. Le seuil de significativité fixé correspond au seuil classique de 5% corrigé pour le nombre de tests indépendants, ici correspondant au nombre de marqueurs testés. Les analyses génétiques ont été effectuées à l'aide du logiciel PLINK v1.90 (Chang *et al.*, 2015; Purcell *et al.*, 2007).

#### ***4.2.4.1 Analyses des données longitudinales***

La médiane et les minimum et maximum de chaque donnée quantitative ont été calculés au moment du diagnostic, et des différents temps de suivi (2, 5 et 10 ans de DT1). Les corrélations entre le taux d'HbA1c à chaque temps ainsi qu'entre la DIR à chaque temps ont été estimées par le coefficient de corrélation de Pearson.

L'évolution du nombre de patients dans chacune des catégories du taux d'HbA1c définies plus haut a été testée à l'aide d'un test du  $\chi^2$ .

Des analyses statistiques plus poussées en prenant en compte le traitement des patients seront effectuées dans un second temps, quand l'ensemble des données aura été collecté pour l'ensemble des patients.

## 4.3 Résultats

Les résultats présentés ici correspondent à l'étude préliminaire des données, la collecte des données n'étant pas terminée.

### 4.3.1 Résultats des données cliniques à 2 ans de diabète

#### 4.3.1.1 Caractéristiques de la cohorte

Les principales données collectées à 2 ans sont résumées dans les tableaux suivants, en fonction du type de données collectées.

Les différents facteurs démographiques collectés dans GENDIR sont présentés dans le Tableau 12. Nous avons collecté les données de 315 patients, dont 47% étaient des garçons, et l'âge médian à 2 ans de DT1 étaient de 9,5 ans. La majorité des patients est issue d'un mariage où les deux parents sont nés en France (64%). La seconde catégorie la plus présente correspond aux enfants dont les deux parents sont nés à l'étranger, et dans 18% des cas ces parents sont nés dans un des trois pays du Maghreb. La CSP était peu fournie dans les dossiers, le métier des parents à l'inclusion étant indiqué dans moins de 40% des dossiers, et la taille de la fratrie est de  $1.9 \pm 1.0$  enfant par famille.

Les différentes caractéristiques des patients au diagnostic sont présentées dans le Tableau 13. La grande variabilité du taux d'HbA1c au moment du diagnostic peut s'expliquer par les différents modes de diagnostic et de la variabilité des symptômes. En effet, certains ont été diagnostiqués suite à un dépistage familial d'apparentés de personnes souffrant de maladie auto-immune dans la famille, alors que d'autres patients arrivent via les urgences suite à une acidocétose diabétique (DKA pour *diabetic ketoacidosis* en anglais) ayant entraîné un coma. 99% des patients ont présenté un syndrome polyuro-polydipsie au moment du diagnostic et 87% une perte de poids.



	N	Médiane [min-max] ou Pourcentage
Genre (% garçon)	315	47%
Age (ans)	309	9.5 [2.5-19.8]
Lieu de naissance du père	269	
France	178	66%
Maghreb	61	23%
Autre	30	11%
Lieu de naissance de la mère	271	
France	198	73%
Maghreb	55	20%
Autre	18	7%
Nombre de parents nés en France	265	
2	169	64%
1	33	12%
0	63	24%
CSP du père	120	
Artisans, commerçants, chefs d'entreprises	12	10%
Cadres et professions intellectuelles supérieures	33	28%
Professions intermédiaires	26	22%
Employés	6	5%
Ouvriers	34	28%
Retraités	1	1%
Autres personnes sans activités professionnelle	8	7%
CSP de la mère	121	
Artisans, commerçants, chefs d'entreprises	2	2%
Cadres et professions intellectuelles supérieures	20	17%
Professions intermédiaires	26	21%
Employés	40	33%
Ouvriers	1	1%
Retraités	0	0%
Autres personnes sans activités professionnelle	32	26%
Taille de la fratrie	240	3 [1-7]

**Tableau 12 :** Facteurs démographiques à deux ans de diabète dans la cohorte GENDIR

	N	Médiane [min-max] ou Pourcentage
Age au diagnostic (ans)	309	7.5 [0.5-18]
HbA1c au diagnostic (%)	240	11.6 [5.7-18.2]
IMC au diagnostic (kg/m <sup>2</sup> )	190	15.4 [10.1-29.8]
Syndrome Polyuro-polydipsie au diagnostic	289	
Oui	285	99%
Non	4	1%
Perte de poids au diagnostic	279	
Oui	242	87%
Non	37	13%
Cétose au diagnostic	248	
Oui	185	75%
Non	63	25%
Acidocétose diabétique au diagnostic	271	
Oui	98	36%
Non	173	64%
Réanimation au diagnostic	233	
Oui	49	21%
Non	184	79%

**Tableau 13 :** Facteurs liés au diagnostic du DT1 dans la cohorte GENDIR

Les différents auto-anticorps (AAC) testés au moment du diagnostic sont présentés dans le Tableau 14.

AAC	N	Pourcentage
Anti-Gad	219	
Positif	169	77%
Négatif	50	23%
Anti-IA2	203	
Positif	154	76%
Négatif	49	24%
Anti-ICA	65	
Positif	41	63%
Négatif	24	37%
Anti-Ins	55	
Positif	41	75%
Négatif	14	25%

**Tableau 14 :** Auto-anticorps (AAC) mesurés au moment du diagnostic de DT1

L'ensemble des données collectées concernant le traitement des patients à 2 ans de DT1 est présenté dans le Tableau 15. À deux ans, seuls 50 patients, soit 20% de la cohorte, sont sous pompe à l'insuline.

Les doses d'insuline varient entre 0.20 et 1.64 U/kg/j, avec une dose médiane de 0.79 U/kg/j et une dose moyenne de 0.80 U/kg/j.

	N	Médiane [min-max] ou Pourcentage
Type de traitement	249	
Injections	199	80%
Pompe	50	20%
DIR (U/kg/j)	307	0.79 [0.20-1.64]

**Tableau 15:** Facteurs liés au traitement du DT1 à deux ans de diabète dans la cohorte GENDIR

Les autres facteurs que nous avons pu collecter sont présentés dans le Tableau 16. L'IMC à 2 ans est compris entre 14 et 30.2 kg/m<sup>2</sup>. Nous avons collecté plusieurs données concernant la naissance des patients. Dans notre cohorte, parmi les patents pour lesquels nous avons l'information, 13% sont prématurés (nés avant 36 semaines d'aménorrhée) et 12% après le terme (nés après 42 semaines d'aménorrhée).

Nous avons aussi collecté le poids et la taille à la naissance.

	N	Médiane [min-max] ou Pourcentage
IMC à 2 ans (kg/m <sup>2</sup> )	266	17.5 [14.0-30.2]
Terme	199	
Prématurité	26	13%
Terme	149	75%
Post-terme	24	12%
Poids à Naissance	247	3.3 [1.1-5.3]
Taille à la Naissance (cm)	231	50 [36-56]

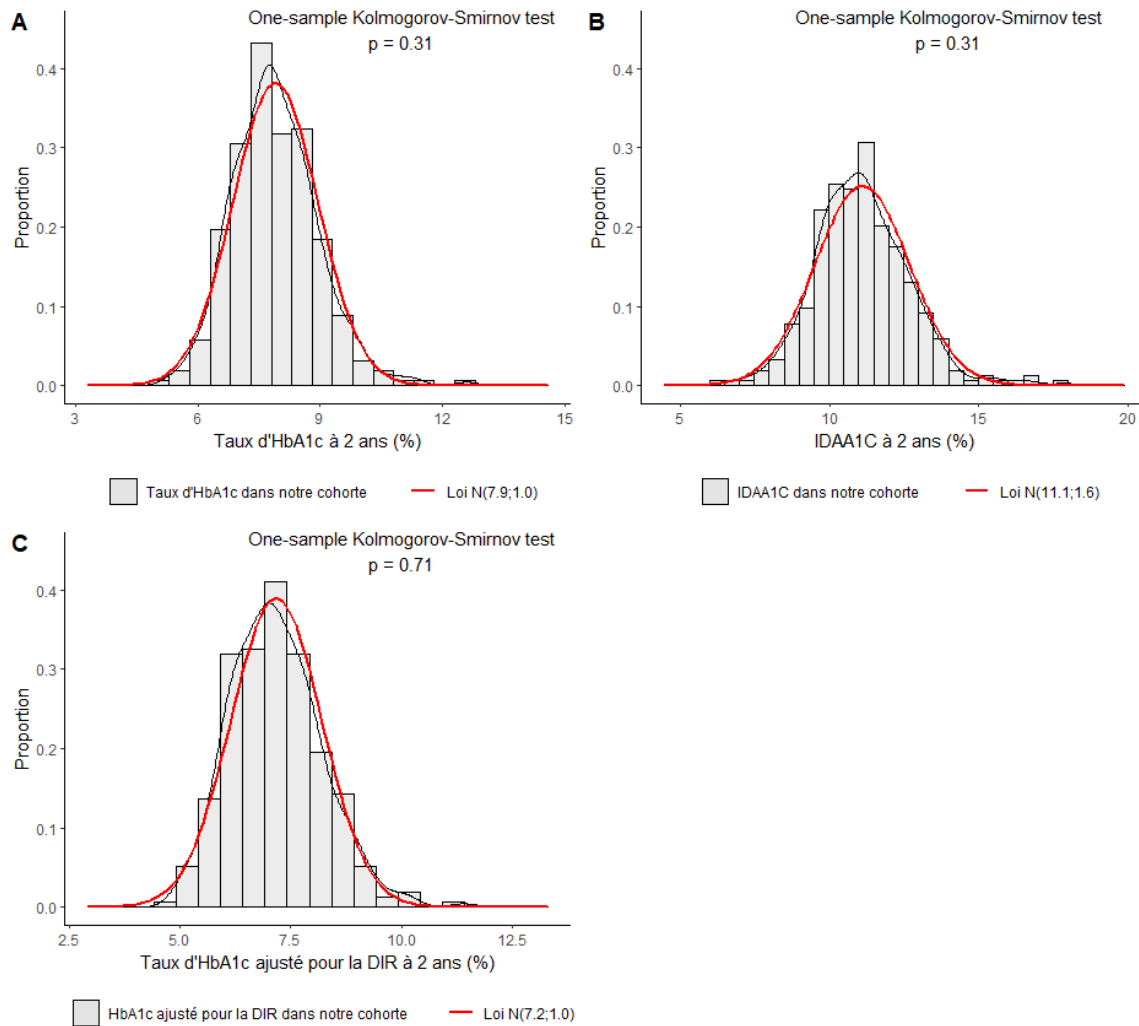
**Tableau 16 :** Autres facteurs collectés dans la cohorte GENDIR

Le taux d'HbA1c, l'IDAA1c et le taux d'HbA1c préajusté pour la DIR à 2 ans sont présentés dans le Tableau 17 et la Figure 17. À 2 ans de DT1, le taux d'HbA1c des patients inclus dans cette étude préliminaire varie entre 5,3% et 12,6% et suit une distribution normale (p=0,31) centrée sur 7,9% et d'écart-type 1. La variable IDAA1C varie entre 6,5 et 17,9 et suit une distribution normale (p=0,31) centrée sur 11.1 et d'écart-type 1,6 ; et enfin l'HbA1c préajusté sur la DIR

varie entre 4,9% et 11,4% et suit une distribution normale ( $p=0,71$ ) centrée sur 7,2% et d'écart-type 1.

	N	Médiane [min-max]
HbA1c (%)	315	7.8 [5.3-12.6]
IDAA1C	307	11.0 [6.5-17.9]
HbA1c préajusté pour la DIR (%)	307	7.1 [4.9-11.4]

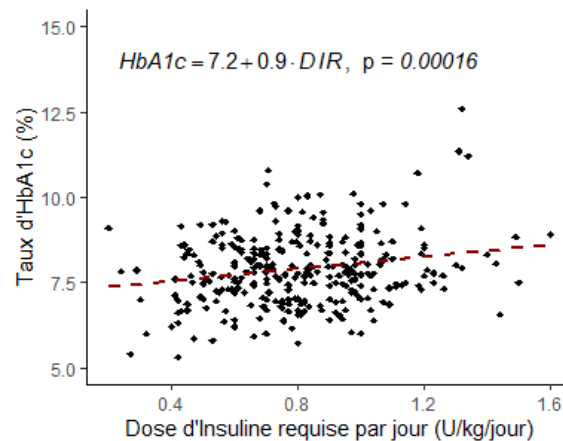
**Tableau 17:** Données glycémiques à 2 ans de DT1



**Figure 17 :** Variation du Taux d'HbA1c (A), de l'IDAA1c (B) et du taux d'HbA1c préajusté pour la DIR (C) à 2 ans et la loi normale associée

#### 4.3.1.2 Analyses des facteurs influençant le taux d'HbA1c après deux ans de DT1

L'effet de la DIR sur le taux d'HbA1c seul est présenté dans la Figure 18.

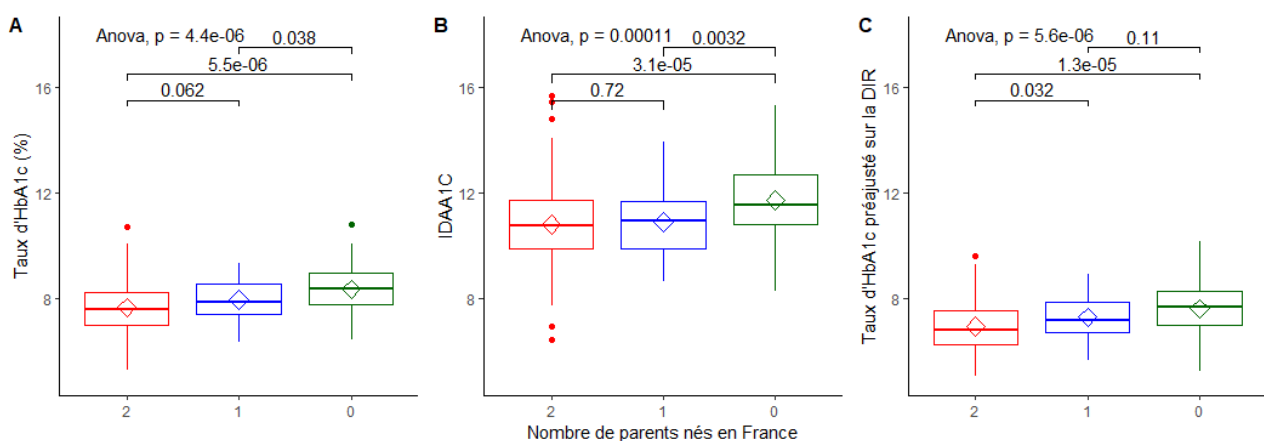


**Figure 18 :** Impact de la DIR sur le taux d'HbA1c à deux ans de diabète

Taux d'HbA1c en fonction de la dose d'insuline requise. Une régression linéaire a été effectuée pour mettre en évidence l'association entre la DIR et le taux d'HbA1c, indiqué en pointillé rouge, l'équation de cette régression est indiquée en haut de la figure, ainsi que le seuil de significativité de la régression.

Notre objectif ici était d'identifier les facteurs influençant le taux d'HbA1c. Nous avons donc corrigé le seuil de signification par le nombre de tests effectués, correspondant ici au nombre de variables testés, soit 14, nous avons ainsi fixé notre seuil à 0.004. Les analyses univariées des facteurs en ajustant pour la DIR à 2 ans de DT1 sont présentées dans le Tableau 18.

L'effet du nombre de parents nés en France est significatif, quelle que soit la méthode d'ajustement du taux d'HbA1c pour la DIR (Figure 19). Dans les trois cas, la différence la plus importante est visible entre le groupe dont les deux parents sont nés en France et le groupe où aucun des deux parents ne sont nés en France, avec un effet intermédiaire pour le groupe ayant un seul parent né en France, laissant penser un effet dose du nombre de parents nés en France.



**Figure 19 :** Impact de l'origine des parents sur le taux d'HbA1c à deux ans de diabète

Taux d'HbA1c (A), IDAA1C (B), et taux d'HbA1c préajusté sur la DIR (C) en fonction du nombre de parents nés en France. Un test d'ANOVA a été fait pour comparer les trois groupes, suivi d'un t-test entre les différents groupes. Le haut et le bas de chaque boîte correspondent au premier et troisième quartile, et la ligne du milieu à la médiane, le losange correspond à la moyenne.

Variables	N	% Données manquantes	Moyenne HbA1c ± SD	P-valeur			
				IDAA1C	HbA1c ajusté	HbA1c préajusté	
<b>Facteurs démographiques</b>							
Genre	315	0%		0.24 *		0.24 *	
	Garçon Fille	148 167		7.9±1.0 7.9±1.0	Ref 0.24 §		
Puberté	309	2%		0.023 μ		0.93 μ	
	Enfant Adolescent Adulte	159 138 12		7.9±1.0 7.9±1.1 7.7±1.5	Ref 0.71 § 0.85 §		
	Nombre de parents nés en France <sup>2</sup>	265	16%		1.1E-4 μ	4.2E-6 μ	
2 1 0	169 33 63		7.7±0.9 8±0.8 8.4±1.0		Ref 0.048 § 5.9E-7 §		
	CSP du père	120	62%		0.56 μ		0.59 μ
		CSP+ CSP- Sans travail	71 40 9		7.7±0.8 7.9±1.0 8.0±1.0	Ref 0.41 § 0.40 §	
CSP de la mère		121	62%		0.0035 μ		0.055 μ
	CSP+ CSP- Sans travail	48 41 32		7.5±0.8 7.8±0.9 8.1±0.9	Ref 0.41 § 7.3E-3 §		
	<b>Facteurs liés au diagnostic du DT1</b>						
Age de début	309	2%	-	0.026 §	0.74 §	0.76 §	
Hba1c au diagnostic	240	24%	-	0.13 §	0.17 §	0.17 §	
Acidocétose diabétique au diagnostic	271	14%		0.0038 *		0.082 *	
	Oui Non	98 173		8.1±1.1 7.8±1.0	Ref 0.066 §		
Cétose au diagnostic	248	21%		0.12 *		0.38 *	
	Oui Non	185 63		7.9±1.0 8.0±1.0	Ref 0.38 §		
Perte de poids au diagnostic	279	11%		0.44 *		0.33 *	
	Oui Non	242 37		7.9±1.1 7.7±0.9	Ref 0.40 §		
Réanimation au diagnostic	233	26%		0.054 *		0.054 *	
	Oui Non	49 184		8.1±1.2 7.8±1.0	Ref 0.034 §		
<b>Facteurs liés au traitement</b>							
Type de traitement	249	21%		1.7E-5 *		0.51 *	
	Injections Pompe	200 50		7.9±1.1 7.6±1	Ref 0.34 §		
<b>Facteurs personnels</b>							
IMC standardisé	260	17%	-	0.54 §	0.011 §	0.011 §	

**Tableau 18** : Seuil de significativité des analyses univariées de l'IDAA1C, du taux d'HbA1c ajusté pour la DIR et du taux d'HbA1c préajusté pour la DIR à 2 ans de DT1

En gras :  $p < 0.004$  ; en italique :  $p < 0.05$  ; \* : testé par le test de l'écart réduit ; § : testé par test de Wald (régression linéaire) ; μ : testé par ANOVA

L'étude de l'IDAA1c montre un effet significatif du statut socio-professionnel de la mère (3.5E-3) et du type de traitement (p=4.1E-5), non significatif en ajustant sur la dose d'insuline.

Les différentes régressions multivariées sont présentées dans le Tableau 19. Avant la régression pas à pas descendante, nous avons testé les modèles suivants :

- Pour l'IDAA1c :

$$IDAA1C = \alpha + \beta_{Ado}Ado + \beta_{Adu}Adu + \beta_{AAO}AAO + \beta_{DKA}DKA + \beta_{Fr}Fr + \beta_P P$$

- Pour l'HbA1c ajusté pendant l'analyse :

$$HbA1c = \alpha + \beta_{DIR}DIR + \beta_{IMC}IMC + \beta_{Fr} Fr$$

- Pour l'HbA1c préajusté :

$$HbA1c_{Préajusté} = \alpha + \beta_{IMC}IMC + \beta_{Fr} Fr$$

Avec :

- *IDAA1C* le phénotype étudié à 2 ans de DT1 correspondant pour chaque patient *i* à  $IDAA1C_i = HbA1c_i + 4 \times DIR_i$  ;
- *HbA1c* le phénotype à 2 ans de DT1 sans ajustement ;
- *HbA1c\_Préajusté* le phénotype correspondant pour chaque patient *i* à  $HbA1c_{Préajusté}_i = HbA1c_i - 0.9 \times DIR_i$  ;
- $\alpha$  l'intercept de la régression et  $\beta_j$  le coefficient de régression de la variable *j* ;
- *Ado* la variable qui vaut 1 si le patient est considéré comme adolescent à 2 ans de diabète et vaut 0 autrement ;
- *Adu* la variable qui vaut 1 si le patient est considéré comme adulte à 2 ans de diabète et vaut 0 autrement ;
- *AAO* la variable correspondant à l'âge de début ;
- *DKA* la variable correspondant à la présence (1) ou l'absence (0) d'acidocétose au diagnostic ;
- *Fr* correspondant à la variable modélisant le nombre de parents nés en France, valant 0, 1 ou 2 en fonction du nombre de parents nés en France ;
- *P* correspondant à l'utilisation d'une pompe à 2 ans de DT1 (1) ou d'un autre type de traitement (0).

- *DIR* correspondant à la valeur de la *DIR* à 2 ans de DT1 ;
- *IMC* correspondant à la valeur de l'*IMC* standardisé.

Au vu des coefficients de régression, nous avons décidé de considérer un effet dose du nombre de parents nés en France (Tableau 19).

	Multivarié		Pas à pas descendante	
	$\beta \pm SE$	p	$\beta \pm SE$	p
<b>IDAA1C</b>				
<b>Intercept</b>	10,85±0,28	<2E-16	10,69±0,17	<2E-16
<b>Puberté</b>				
<b>Enfant</b>	Ref			
<b>Adolescent (Ado)</b>	0,74±0,37	<b>0,044</b>	0,41±0,21	<b>0,046</b>
<b>Adulte (Adu)</b>	1,09±0,69	0,12	0,52±0,48	0,28
<b>Age de début (AAO)</b>	-0,053±0,045	0,25	-	-
<b>Nombre de parents nés en France (Fr)</b>	0,41±0,13	<b>0,0013</b>	0,45±0,13	<b>0,00049</b>
<b>Acidocétose diabétique au diagnostic (DKA)</b>				
<b>Non</b>	Ref		-	-
<b>Oui</b>	0,30±0,21	0,14	-	-
<b>Type de traitement (P)</b>				
<b>Injections</b>	Ref			
<b>Pompe</b>	-0,77±0,25	<b>0,002</b>	-0,67±0,24	<b>0,0052</b>
<b>HbA1c Ajusté</b>				
<b>Intercept</b>	7,22±0,21	<2E-16	7,27±0,21	<2E-16
<b>DIR (DIR)</b>	0,60±0,25	0,019	0,56±0,25	<b>0,029</b>
<b>IMC standardisé (IMC)</b>	0,095±0,068	<b>0,016</b>	-	-
<b>Nombre de parents nés en France (Fr)</b>	0,27±0,075	<b>0,00036</b>	0,28±0,075	<b>0,00021</b>
<b>HbA1c Préajusté</b>				
<b>Intercept</b>	6,98±0,073	<2E-16	7,0±0,073	<2E-16
<b>IMC standardisé (IMC)</b>	0,10±0,068	0,12	-	-
<b>Nombre de parents nés en France (Fr)</b>	0,27±0,075	<b>0,00043</b>	0,28±0,075	<b>0,00024</b>

**Tableau 19** : Résultats des régressions multivariées et pas à pas descendantes

Les régressions pas à pas descendantes pour l'analyse du taux d'HbA1c ajusté pendant l'analyse ou préajusté retiennent une seule variable : le nombre de parents nés en France. Au vu des résultats en uni et multivarié, les résultats en utilisant le taux d'HbA1c corrigé préajusté ou ajusté au cours de l'analyse sont très proches. Nous avons donc décidé d'utiliser le taux d'HbA1c préajusté sur la *DIR* pour la partie génétique, afin de ne pas sous-évaluer l'effet de la *DIR* dans le sous-échantillon génotypé.

En revanche, l'*IDAA1C* ne donne pas des résultats similaires, puisque l'analyse pas à pas descendante permet de retenir en plus de l'effet du nombre de parents nés en France celui de



la puberté et du type de traitement (Tableau 19). L'association de ces deux facteurs avec l'IDAA1C peut s'expliquer par le fait que ces deux facteurs sont significativement associés avec la DIR ( $p=9.6E-5$  pour le fait d'être adolescent et  $p=3.6E-4$  pour le fait d'être sous pompe). L'IDAA1C n'a pas été développée pour étudier le taux d'HbA1c sur le long terme, mais pour prédire l'évolution du peptide-C pendant la rémission partielle, soit avant les 2 ans de DT1. Au vu de ces résultats et de la construction de l'IDAA1c, nous avons décidé de ne pas l'étudier pour la partie génétique puisque l'ajustement semble augmenter le poids de la DIR. Cela demande des analyses complémentaires, qui seront faites sur l'ensemble des données.

### 4.3.2 Analyse des données génétiques

#### 4.3.2.1 Description des données génétiques

L'ensemble de la cohorte GENDIR n'a pas été génotypé et l'ensemble des données cliniques n'a pas été saisie pendant ma thèse pour les patients génotypés.

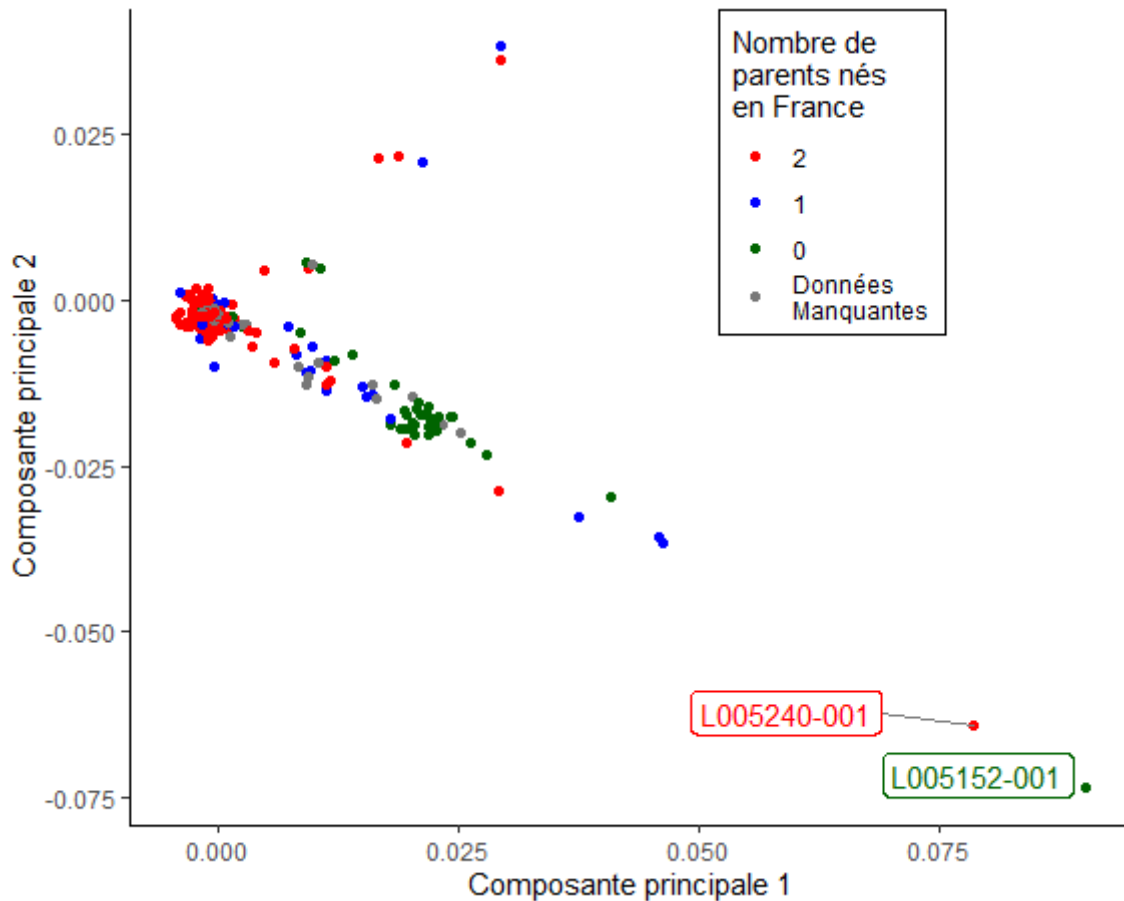
Sur les 260 patients génotypés, 16 patients ont été exclus suite au contrôle qualité des données génétiques sur les critères suivants :

- Plus de 5% de génotypage manquant ;
- Sexe clinique et génotypique différents ;
- Taux d'hétérozygotie non compris entre 0,29 et 0,35 (correspondant aux bornes du taux d'hétérozygotie  $\pm 3$  écarts types).

Deux patients ont été exclus au moment de la vérification de la stratification de la population à l'aide de la PCA, puisque trop éloignés des autres patients (L005240-001 et L005152-001). Ces deux *outliers* sont indiqués dans la Figure 20. Les parents du patient L005240-001 sont nés aux Comores et en Angola, et les parents du patient L005152-001 sont tous les deux nés en France. Cependant l'information quant au lieu de naissance des parents ne nous informe pas totalement sur l'origine ethnique, ce qui peut expliquer le cas du patient L005152-001.

Cette figure montre aussi une corrélation entre les deux premières composantes principales et le nombre de parents nés en France (corrélation de Pearson de  $r=0.62$ ,  $p<2E-16$  entre la composante 1 et le nombre de parents nés en France, et  $r=-0.49$ ,  $p=6.3E-13$  entre la composante 2 et le nombre de parents nés en France), avec un cluster de patients

correspondant aux patients ayant deux parents nés en France, et un second correspondant aux patients dont les deux parents ne sont pas nés en France. En raison de cette relation, nous avons choisi d'inclure les deux premières composantes dans les analyses génétiques et pas la variable nombre de parents nés en France, qui n'était pas toujours renseignée.



De plus, il y a 35 patients génotypés pour lesquels nous n'avons pas encore les données du taux d'HbA1c à 2 ans de diabète ou les données de la dose d'insuline. Les différentes variables collectées sont globalement similaires entre les 211 patients génotypés et l'échantillon de départ de 315 patients, comme nous pouvons le voir dans le Tableau 20.

	GENDIR		Patients génotypés	
	N	Médiane [min-max] ou Pourcentage	N	Médiane [min-max] ou Pourcentage
<b>Facteurs démographiques</b>				
<b>Genre (% garçon)</b>	315	47%	211	45%
<b>Age (ans)</b>	309	9.5 [2.5-19.8]	209	9.2 [2.5-17.9]
<b>Nombre de parents nés en France</b>	265		189	
<b>2</b>	169	64%	131	69%
<b>1</b>	33	12%	26	14%
<b>0</b>	63	24%	32	17%
<b>Facteurs liés au traitement</b>				
<b>Type de traitement</b>	249		202	
<b>Injections</b>	199	80%	158	78%
<b>Pompe</b>	50	20%	44	22%
<b>DIR (U/kg/j)</b>	307	0.79 [0.20-1.64]	211	0.78 [0.20 – 1.64]
<b>Facteurs personnels</b>				
<b>IMC à 2 ans (kg/m<sup>2</sup>)</b>	266	17.5 [14.0-30.2]	189	17.2 [14.0-30.5]
<b>Données glycémiq</b>				
<b>HbA1c (%)</b>	315	7.8 [5.3-12.6]	211	7.8 [5.3-10.8]
<b>IDAA1C (%)</b>	307	11.0 [6.5-17.9]	211	10.9 [6.5-15.7]
<b>HbA1c préajusté pour la DIR (%)</b>	307	7.1 [4.9-11.4]	211	7.1 [4.9-10.2]

**Tableau 20** : Descriptif de la cohorte GENIR à 2 ans et du sous-échantillon génotypé

#### 4.3.2.2 Etude de puissance et choix des marqueurs

Concernant les marqueurs identifiés dans les populations non-diabétiques comme étant associés au taux d'HbA1c (présentés dans le Chapitre 1.3.4.2), ceux-ci ont des effets reportés dans GWAS Catalog inférieur à 0.1% de la variance expliquée, ce qui ne sera pas détectable dans notre échantillon de 211 patients. Nous avons donc décidé de ne pas les étudier.

Trente marqueurs ont été reportés comme étant significativement associés dans au moins une étude du taux d'HbA1c chez les patients DT1, présents dans le Tableau 21 et en Annexe 5, classés de la même manière que dans notre revue (Gloaguen *et al.*, 2018) : quatre gènes ont été retrouvés associés à au moins une maladie auto-immune ; neuf gènes sont impliqués dans l'immunité ; six sont impliqués dans le métabolisme et aucun dans le cycle de vie des érythrocytes. Parmi ces 30 marqueurs, dix sont présents sur la puce et ont passé le contrôle qualité.

Chr	Gène d'intérêt	Classification	Variant	MAF	Marqueurs en LD	LD (r <sup>2</sup> )	Publication
1	PTPN22	Maladie auto-immune ; Immunité	rs2476601	0.06	rs6679677	0.97	(Petroni <i>et al.</i> , 2008)
	LPPR4	Métabolisme	rs7514675	0.23	/	/	(Syreeni <i>et al.</i> , 2019)
2	CTLA4	Maladie auto-immune ; Immunité	rs5742909	0.052	/	/	(Almasi <i>et al.</i> , 2015)
3	CCR5	Métabolisme; Immunité	rs333	0.072	rs6808835	0.74	(Slominski <i>et al.</i> , 2019)
	RAF1	Metabolism	rs144918527	0.561	/	/	(Syreeni <i>et al.</i> , 2019)
4	TLR3	Immunité	rs11721827	0.113	/	/	(Assmann <i>et al.</i> , 2014)
			rs5743313	0.186	/	/	
5	FLJ23577	Autre	rs286405	0.36	rs286411	0.99	(Paterson <i>et al.</i> , 2010)
6	MEAT6	Autre	rs373883321	0.36	/	/	(Syreeni <i>et al.</i> , 2019)
	TNF	Maladie Auto-immune ; Immunité	rs361525	0.061	rs3093727	0.71	(Korolija <i>et al.</i> , 2010)
7	DPP6	Autre	rs478414	0.44	/	/	(Syreeni <i>et al.</i> , 2019)
	IL6	Immunité	rs1800795	0.14	rs1800795	1	(Urahy <i>et al.</i> , 2015)
8	TMEM74	Autre	rs113241624	0.001	/	/	(Syreeni <i>et al.</i> , 2019)
9	BNC2	Métabolisme; Immunité	rs10810632	0.197	rs4961760	0.98	(Paterson <i>et al.</i> , 2010)
			rs6475082	0.21		0.98	
			rs4961760	0.146		1	
	SYK	Immunité	rs4744017	0.39	/	/	(Syreeni <i>et al.</i> , 2019)
10	SORCS1	Autre	rs1358030	0.478	/	/	(Paterson <i>et al.</i> , 2010)
11	LRP5	Métabolisme	rs3736228	0.14	rs3736228	1	(Souza <i>et al.</i> , 2018)
12	ERBB3	Maladie Auto-immune ; Métabolisme; Immunité	rs2292239	0.3	rs2292239	1	(Brorsson <i>et al.</i> , 2016)
13	RXFP2	Autre	rs2085277	0.105	/	/	(Syreeni <i>et al.</i> , 2019)
15	FANCI	Autre	rs552894079	0.004	/	/	(Syreeni <i>et al.</i> , 2019)
	WDR72	Autre	rs493218	0.223	rs690399	1	(Paterson <i>et al.</i> , 2010)
			rs572221	0.234		0.98	
			rs690271	0.234		0.98	
			rs566369	0.268		0.98	
rs482541	0.268	0.98					
18	C18orf16	Autre	rs163061	0.368	/	/	(Paterson <i>et al.</i> , 2010)
19	PHLDB3	Autre	rs2599441	0.178	rs11668793	0.99	(Syreeni <i>et al.</i> , 2019)

**Tableau 21 :** Marqueurs identifiés comme étant associés au taux d'HbA1c dans les différentes publications, publiées avant septembre 2019  
MAF rapportée dans dbSNP par Exac Global ou 1000 Génome.

Nous avons effectué une étude de puissance avec comme paramètres :

- Une taille de la population de 211 individus ;
- Une moyenne du trait correspondant à la moyenne du taux de l'HbA1c de  $7.9 \pm 1.0\%$  ;
- Un seuil de significativité de 0.005, correspondant au seuil classique de 0.05 sur lequel nous avons appliqué la correction de Bonferroni, soit 0.05 divisé par le nombre de marqueurs testés (10).

Nous obtenons ainsi une puissance de 80% de détecter un variant avec une MAF de 5% et un pourcentage de variance expliqué de 0.62% pour les données du taux d'HbA1c disponibles.

#### 4.3.2.1 Analyse d'association

Afin de tester l'association entre le taux d'HbA1c et les marqueurs génétiques sélectionnés, nous avons décidé d'utiliser le taux d'HbA1c préajusté sur la DIR calculé à partir des données de la cohorte totale de 315 patients. Nous obtenons donc finalement le modèle suivant pour chaque marqueur :

$$HbA1c\_Préajusté = \alpha + \beta_{Add}Add + \beta_{DomDev}DomDev + \beta_{PC1}PC1 + \beta_{PC2}PC2 + E$$

Avec :

- *HbA1c\_Préajusté* le phénotype correspondant pour chaque patient  $i$  à  $HbA1c\_Préajusté_i = HbA1c_i - 0.9 \times DIR_i$  ;
- *Add* la variable qui permet de modéliser l'effet dose de l'allèle à risque, codé 0, 1 et 2 pour respectivement 0, 1 ou 2 allèles à risque ;
- *DomDev* la variable qui modélise l'écart à l'additivité codée 0, 1 et 0 ;
- PC1 et PC2 les deux premières composantes de l'analyse en composante principale.

Les résultats de l'analyse d'association, correspondant pour chaque variable au coefficient de régression de la variable, son erreur type et sa p-valeur, de ces 10 marqueurs sont présentés dans le Tableau 22.

Gène	Marqueurs	$\beta_{Add} \pm SE$	$p_{Add}$	$\beta_{DomDev} \pm SE$	$p_{DomDev}$	$\beta_{PC1} \pm SE$	$p_{PC1}$	$\beta_{PC2} \pm SE$	$p_{PC2}$	$P_{GRM}$
<b>PTPN22</b>	rs6679677	-0.0067 ± 0.12	0.95	-0.33 ± 0.16	0.037	14.72 ± 8.18	0.073	-11.24 ± 9.15	0.22	0.056
<b>CCR5</b>	rs6808835	-0.49 ± 0.34	0.16	0.90 ± 0.37	0.016	13.77 ± 8.18	0.094	-15.13 ± 9.02	0.095	0.013
<b>FLJ23577</b>	rs286411	0.14 ± 0.14	0.32	-0.20 ± 0.18	0.26	11.05 ± 8.38	0.19	-13.64 ± 9.11	0.14	0.49
<b>TNF</b>	rs3093726	-0.050 ± 0.096	0.61	-0.13 ± 0.14	0.34	15.63 ± 8.29	0.061	-13.84 ± 9.34	0.14	0.53
<b>IL6</b>	rs1800795	-0.18 ± 0.18	0.30	0.0064 ± 0.21	0.98	14.28 ± 8.21	0.083	-15.09 ± 9.20	0.10	0.32
<b>BNC2</b>	rs4961760	0.062 ± 0.15	0.67	-0.19 ± 0.18	0.32	14.79 ± 8.29	0.76	-12.97 ± 9.20	0.16	0.58
<b>LRP5</b>	rs3736228	0.031 ± 0.098	0.75	0.22 ± 0.14	0.11	15.27 ± 8.25	0.066	-13.62 ± 9.12	0.14	0.24
<b>ERBB3</b>	rs2292239	0.041 ± 0.12	0.72	0.045 ± 0.15	0.77	15.21 ± 8.25	0.067	-13.42 ± 9.23	0.15	0.82
<b>WDR72</b>	rs690399	-0.37 ± 0.48	0.44	0.51 ± 0.53	0.34	14.99 ± 8.25	0.071	-14.65 ± 9.20	0.11	0.61
<b>PHLDB3</b>	rs11668793	-0.0022 ± 0.091	0.98	-0.022 ± 0.14	0.88	14.91 ± 8.26	0.073	13.88 ± 9.20	0.13	0.99

**Tableau 22** : Analyse des marqueurs associés au taux d'HbA1c préajusté pour la DIR chez les patients DT1 dans notre échantillon

Aucun marqueur n'est significativement associé au taux d'HbA1c dans notre cohorte pour cette étude préliminaire. Le marqueur rs6808835 dans le gène CCR5 semble toutefois suggéré au seuil de 5%. Le modèle retenu par le GRM pour ce SNP correspond à un mode de transmission récessif pour l'allèle C, qui est le plus fréquent dans la population, ce qui est cohérent avec le mode de transmission de la délétion  $\Delta 32$  (Slominski *et al.*, 2019). Ce marqueur est présent dans le gène CCRL2 (*CC chemokine receptor-like 2*) et est en déséquilibre de liaison avec le marqueur rs333, lui-même dans le gène CCR5 (*C-C chemokine receptor type 5*). Tous deux sont des gènes impliqués dans la réponse immunitaire.

### 4.3.3 Analyses des profils d'évolution d'équilibre glycémique

L'objectif de cette partie est d'identifier les différents profils d'évolution du taux d'HbA1c, afin dans un second temps, de déterminer si des profils d'évolution particuliers peuvent être associés à des facteurs spécifiques, génétiques ou non.

#### 4.3.3.1 Descriptif de la cohorte à chaque temps

L'évolution des différents paramètres et du taux d'HbA1c sont présentées dans le Tableau 23. Les patients de la cohorte ont été diagnostiqués entre 1990 et 2009 et n'avaient pas tous des données de suivi à 10 ans disponibles, quand la collecte des données a débuté. Un certain nombre de patients sont aussi considérés à 5 ou 10 ans comme perdus de vue, notamment au moment du passage de la diabétologie pédiatrique à celle adulte.

	Diagnostic		2 ans		5 ans		10 ans	
	N	Médiane [min-max] ou Pourcentage	N	Médiane [min-max] ou Pourcentage	N	Médiane [min-max] ou Pourcentage	N	Médiane [min-max] ou Pourcentage
<b>Facteurs démographiques</b>								
Age (ans)	248	7.3 [0.5-16.3]	309	9.5 [2.5-19.8]	271	12.1 [5.5-22.8]	148	15.6 [10.8-25.9]
<b>Facteurs liés au traitement</b>								
Type de traitement	197		249		228		146	
Injections	89%		80%		62%		46%	
Pompe	11%		20%		38%		54%	
DIR (U/kg/j)	-	-	307	0.79 [0.20-1.64]	258	0.90 [0.37-1.93]	145	0.87 [0.40-1.79]
<b>Facteurs personnels</b>								
IMC (kg/m <sup>2</sup> )	182	15.2 [10.5-29.8]	266	17.5 [14.0-30.2]	250	18.7 [13.2-34.1]	144	20.7 [12.9-35.6]
<b>Données glycémiq</b>								
HbA1c (%)	251	11.5 [5.7-18.2]	315	7.8 [5.3-12.6]	277	7.9 [5.5-12.6]	149	8.1 [5.8-13.1]
IDAA1C	-	-	307	11.0 [6.5-17.9]	258	11.5 [7.1-18.3]	145	11.7 [8.1-17.2]

**Tableau 23** : Évolution des différentes variables au cours du suivi

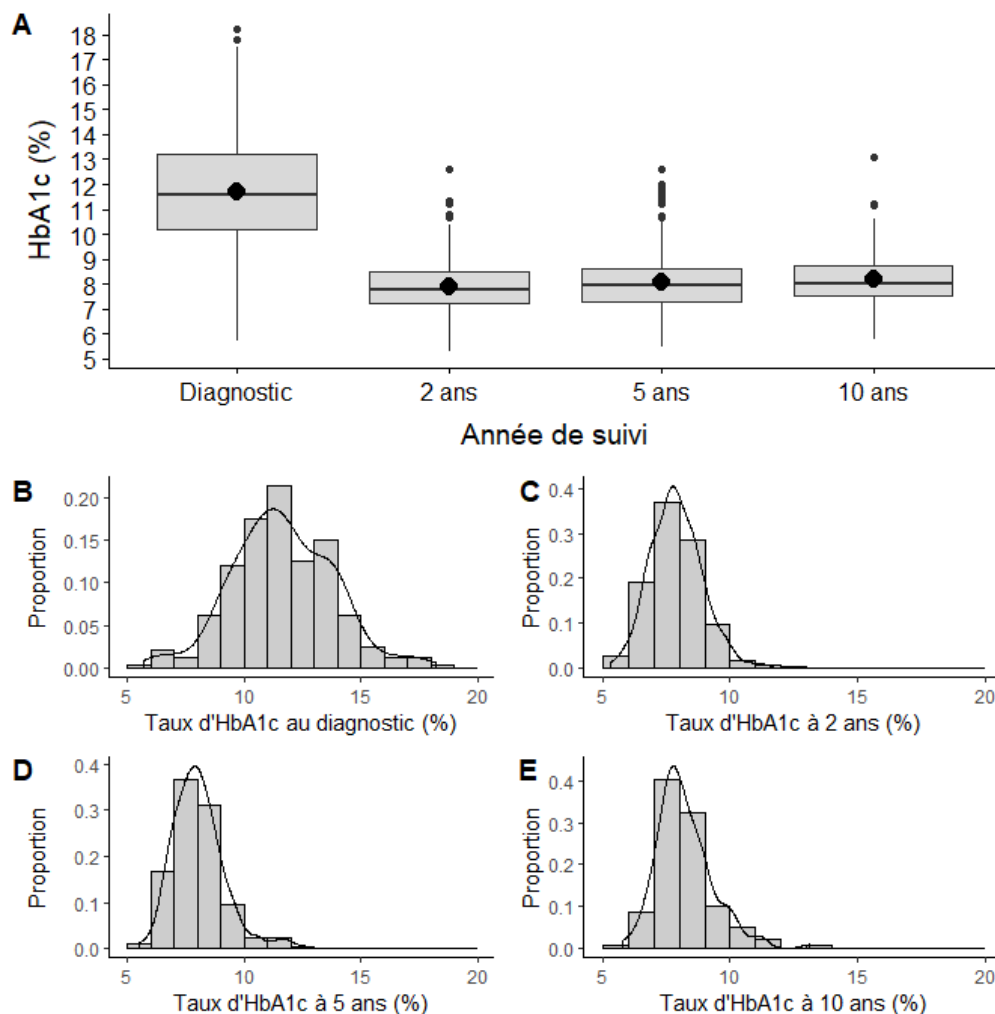
Il existe une augmentation de la proportion de patients sous pompe, passant de 20% à 2 ans de DT1, à 54% après 10 ans de DT1. Ceci peut refléter en partie le développement ces dernières années de la pompe à insuline, mais aussi par le fait que l'âge des patients évoluent et il est connu que l'utilisation et l'adhérence à la pompe sont fortement associées à l'âge du patient (Miller *et al.*, 2015; Spaans, Kleefstra, Groenier, Bilo, & Brand, 2020). De plus, la construction de notre cohorte avec des dates de diagnostic étalées sur 20 ans peuvent rendre complexe l'exploitation de cette donnée. Une étude plus précise sera effectuée quand l'ensemble des données sera collecté.

Les taux d'HbA1c aux différents temps de suivi sont présentés dans la Figure 21. À 2, 5 et 10 ans, la variabilité du taux d'HbA1c est faible, avec une moyenne de 7.9% (comprise entre 5.3% et 12.6%), 8.0% (comprise entre 5.2% et 12.6%) et 8.2% (comprise entre 5.8% et 13.1%) à respectivement 2, 5 et 10 ans.

L'étude de l'évolution du taux d'HbA1c moyen durant le suivi a montré que le taux d'HbA1c à 5 ans ainsi que le taux d'HbA1c à 10 ans sont corrélés au taux d'HbA1c à 2 ans, avec respectivement un coefficient de corrélation de 0.44 et de 0.40 (Figure 22). Le taux d'HbA1c à 5 ans est aussi corrélé au taux d'HbA1c à 10 ans ( $r=0.48$ ). En revanche, aucune corrélation n'est



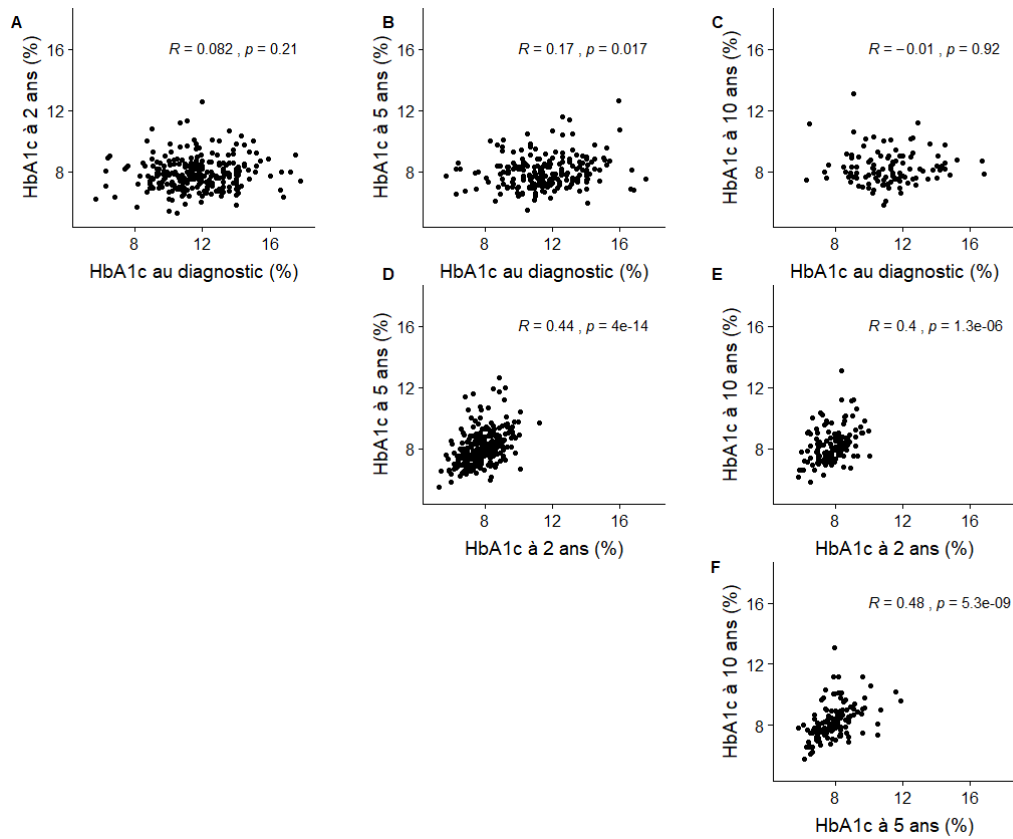
identifiée entre le taux d'HbA1c au diagnostic et le taux d'HbA1c à 2 et 10 ans, et une faible corrélation entre le taux d'HbA1c au diagnostic et à 5 ans.



**Figure 21 :** Valeur du taux d'HbA1c au diagnostic, à 2, 5 et 10 ans dans notre cohorte

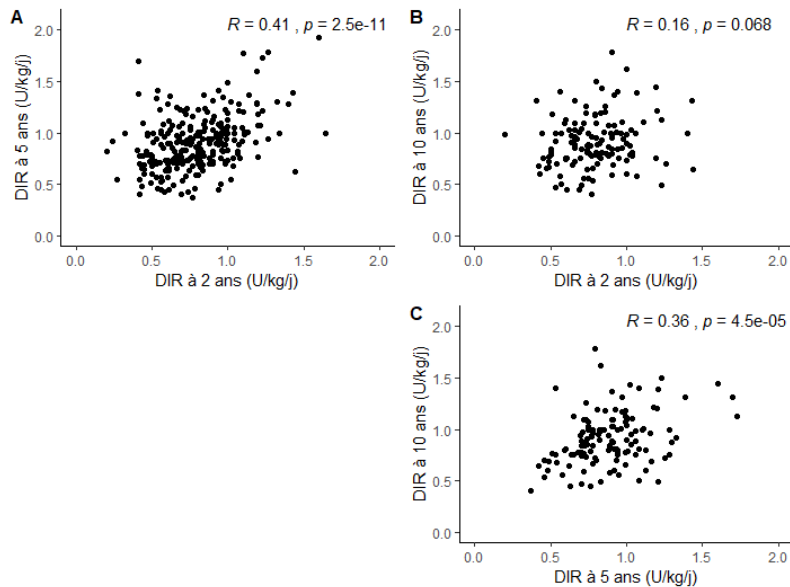
A- Le taux d'HbA1c au moment du diagnostic, à 2, 5 et 10 ans. B- Histogramme du taux d'HbA1c au diagnostic, C- à 2 ans, D- à 5 ans et E- à 10 ans.

De la même manière, nous avons regardé la corrélation entre la dose d'insuline requise à chaque temps (Figure 23). La DIR est corrélée entre 2 ans et 5 ans, ainsi qu'entre 5 ans et 10 ans, avec un coefficient de corrélation de respectivement de 0.41 et de 0.36. Entre 2 ans et 10 ans, la corrélation n'est pas significative ( $p=0.068$ ).



**Figure 22 :** Nuage de points du taux d'HbA1c aux différents temps d'analyse et le coefficient de corrélation de Pearson correspondant

A- Entre le taux d'HbA1c au diagnostic et à 2 ans, B- entre le taux d'HbA1c au diagnostic et à 5 ans, C- entre le taux d'HbA1c au diagnostic et à 10 ans, D- entre le taux d'HbA1c à 2 ans et à 5 ans, E- entre le taux d'HbA1c à 2 ans et à 10 ans, F- entre le taux d'HbA1c à 5 ans et à 10 ans



**Figure 23 :** Nuage de points de la DIR aux différents temps d'analyse et le coefficient de corrélation de Pearson correspondant  
A- entre la DIR à 2 ans et à 5 ans ; B- entre la DIR à 2 ans et à 10 ans ; C- entre la DIR à 5 ans et à 10 ans

La présence de complications ainsi que de différentes comorbidités durant le suivi, telles que les allergies et d'autres maladies auto-immunes, ont été collectées et sont présentées dans le Tableau 24.

	N	Médiane [min-max] ou Pourcentage
Complication à 10 ans	315	
Oui	6	2%
Non	309	98%
Allergie à 10 ans	159	
Oui	61	38%
Non	98	62%
Autre maladie auto-immune à 10 ans	276	
Oui	69	25%
Non	207	75%

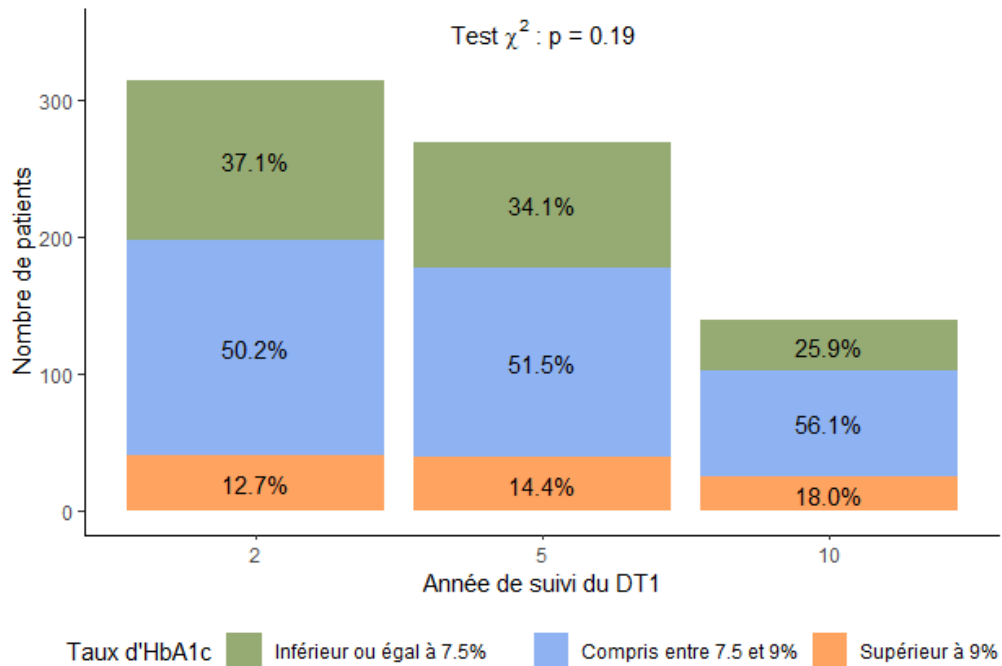
**Tableau 24 :** Développement des complications et des comorbidités après 10 ans de maladie

Nous pouvons noter la présence de six complications après 10 ans de DT1 (cinq rétinopathies et une néphropathie), chez six filles. Concernant les allergies, 38% des patients ont développé durant le suivi au moins une allergie, et 25% ont développé une autre maladie auto-immune, et notamment des thyroïdites d'Hashimoto.

#### *4.3.3.2 Evolution du taux d'HbA1c au cours du temps*

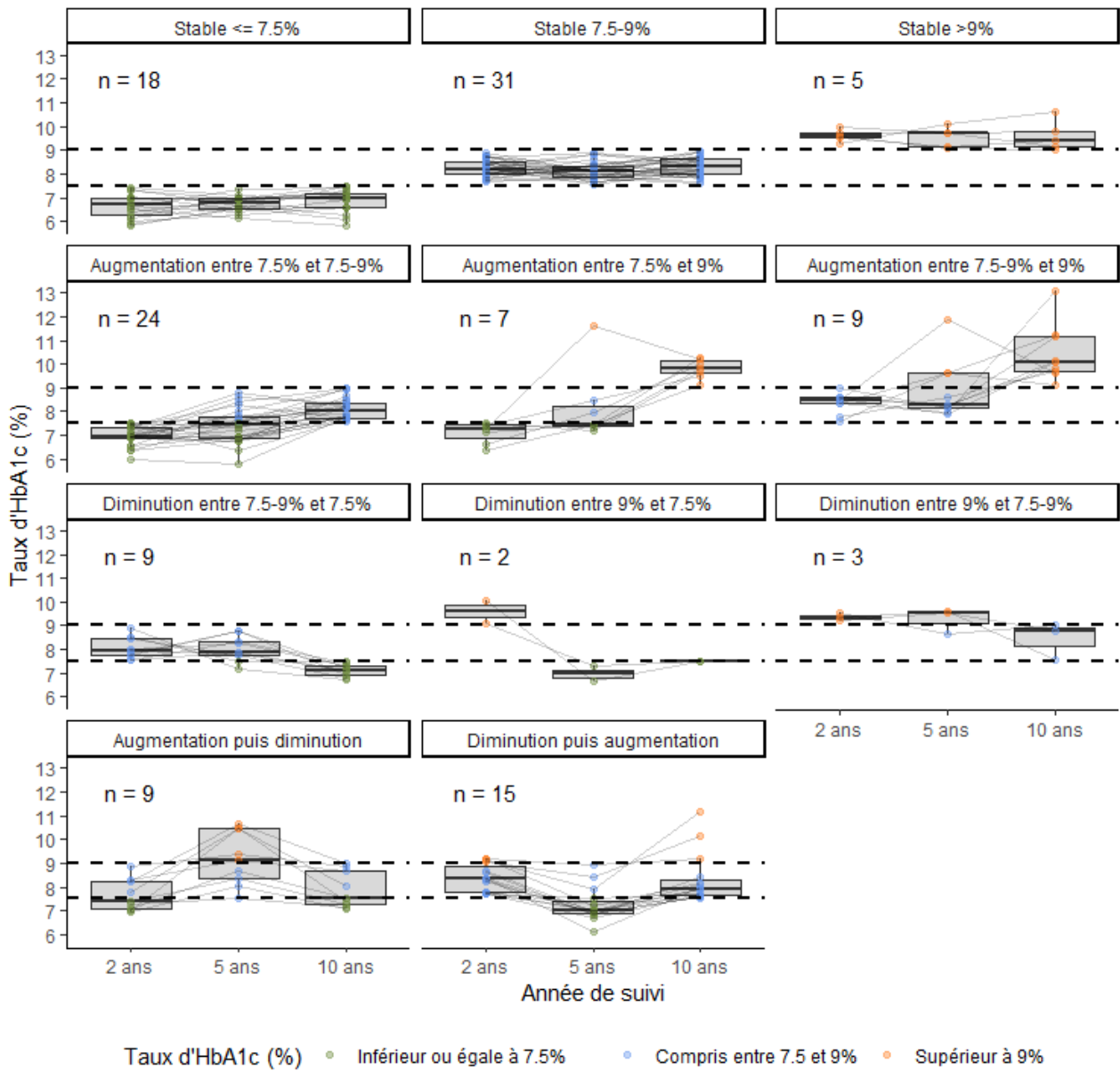
Dans notre cohorte, à deux ans de diabète, 37.1% des patients ont un taux d'HbA1c inférieur à 7.5%, considéré comme un bon objectif par les cliniciens, et à l'inverse 12.7% ont un taux d'HbA1c supérieur à 9% (Figure 24). Durant l'évolution du DT1, ces deux taux varient et passent respectivement à 34.1% et 14.4% à 5 ans et 25.9% et 18% à 10 ans de DT1. Cependant, cette différence n'est pas significative ( $p=0.19$ ).

À partir des données au diagnostic et à 2, 5 et 10 ans, il est possible de définir des profils d'évolution du taux d'HbA1c. Nous avons ainsi pu représenter les différents profils pour les 132 patients pour lesquels nous avons l'ensemble des données du taux d'HbA1c à 2, 5 et 10 ans (Figure 25).



**Figure 24 :** Évolution du nombre de patients en fonction de la catégorie du taux d'HbA1c

Dix-huit patients ont toujours un taux d'HbA1c inférieur à 7.5% et cinq toujours un taux d'HbA1c supérieur ou égale à 9%. La majorité des patients ayant un taux d'HbA1c stable, c'est-à-dire qui reste dans une même catégorie, ont un taux d'HbA1c toujours compris entre ces deux valeurs (31 patients). Quarante patients ont un taux d'HbA1c qui augmente entre leurs 2 ans de DT1 et leurs 10 ans, et quatorze ont leur taux d'HbA1c qui diminue entre leurs 2 et 10 ans de DT1. Vingt-quatre autres patients ont des profils d'évolution plus complexes avec une augmentation puis une diminution pour 9 patients, et une diminution puis une augmentation pour 15.



**Figure 25 :** Évolution du taux d'HbA1c des 132 patients qui possèdent des données à 2, 5 et 10 ans  
Le n représente le nombre d'individus pour chaque profil

## 4.4 Discussion

Cette étude préliminaire avait comme objectif principal de rechercher les principaux facteurs pouvant influencer le taux d'HbA1c dans notre cohorte française. En plus de l'impact de la dose d'insuline requise, nous retrouvons l'effet important de l'origine géographique des parents. Au vu de la taille de notre échantillon, ces résultats sont cohérents avec ceux présentés dans notre revue (Gloaguen *et al.*, 2018). D'un point de vue génétique, aucun marqueur n'a été retrouvé comme significativement associé au taux d'HbA1c, même si nous pouvons noter qu'un marqueur proche de la délétion  $\Delta 32$  est suggéré avec un mode de transmission génétique identique à celui présenté dans l'étude de Slominski *et al* (Slominski *et al.*, 2019).

Afin d'étudier l'impact des facteurs que nous avons collectés en corrigeant pour la dose d'insuline requise, nous avons décidé de comparer différentes méthodes pour prendre en compte la DIR. Nous avons effectué trois analyses en parallèle prenant en compte la DIR de façon différente : une analyse en prenant en compte la DIR comme covariable dans le modèle de régression, une analyse en préajustant le taux d'HbA1c par la DIR et enfin en utilisant l'IDAA1c.

De façon attendue, la prise en compte de la DIR pendant ou avant donne des résultats identiques quant à l'effet fort de l'origine des parents. Nous avons donc pour la partie génétique décidé d'utiliser le taux d'HbA1c préajusté. Ce choix a été fait puisque nous ne possédons pas encore l'ensemble des données pour les patients qui ont été génotypés, il sera sans doute plus intéressant au moment de l'analyse final de prendre la DIR directement dans le modèle comme une covariable. En effet, le pré-ajustement du taux d'HbA1c force à fixer l'effet de la DIR, alors qu'il pourrait être intéressant d'étudier les interactions qui pourraient exister entre la part génétique du taux d'HbA1c et la dose d'insuline.

La troisième façon d'étudier le taux d'HbA1c que nous avons utilisé correspond à l'IDAA1c. Un plus grand nombre de facteurs ont été identifiés comme étant significativement associés à l'IDAA1c par rapport à l'étude du taux d'HbA1c ajusté pour la DIR pendant ou avant analyse. L'analyse multivariée de l'IDAA1C fait ressortir en plus de l'effet du nombre de parents nés en France, l'effet du type de traitement et du statut adolescent, deux variables fortement associées à la DIR.

Il est important de noter que l'IDAA1C a été développée pour étudier la rémission partielle, avec une bonne corrélation entre l'IDAA1C et le taux de peptide-C (Mortensen *et al.*, 2009). L'IDAA1c est donc essentiellement étudiée pour la ou les deux premières années de DT1 et à l'aide des deux catégories permettant de définir la rémission partielle : les patients ayant un IDAA1C inférieur ou égale à 9 seraient en rémission partielle versus les patients ayant un IDAA1C supérieur à 9 (Chiavaroli *et al.*, 2019; Hao *et al.*, 2016; Nagl *et al.*, 2017; Redondo *et al.*, 2018). Les patients participant aux études qui ont proposé ou validé l'IDAA1c, avaient au moment du diagnostic un âge équivalent à celui de notre échantillon à deux ans de DT1, alors que nous nous attendions à ce qu'ils soient plus jeunes, ces études étant réalisées au moment du diagnostic (Max Andersen *et al.*, 2014; Mortensen *et al.*, 2009). Les patients de notre échantillon ont donc en moyenne un âge de diagnostic de deux ans plus jeune que celles qui ont permis de développer l'IDAA1c, ce qui pourrait influencer son utilisation sachant qu'il y a une forte corrélation entre le statut pubertaire et la DIR. De manière plus générale, cela pose la question de l'utilisation de l'IDAA1C, ne prenant pas en compte le statut pubertaire, pour lequel l'effet n'est pas linéaire (Miller, Beck, Foster, & Maahs, 2020; Miller *et al.*, 2015).

Le principal facteur identifié par cette étude préliminaire est le nombre de parents nés en France, reflétant en partie l'effet de l'ethnie dans notre cohorte, facteur qui avait été précédemment reporté dans 12 études que nous avons référencées dans notre revue (Cadario *et al.*, 2014; Campbell *et al.*, 2014; Coulon *et al.*, 2017; Fredheim *et al.*, 2013; Hilliard *et al.*, 2013; Hilmi *et al.*, 2013; Kim, Elmi, Henderson, Cogen, & Kaplowitz, 2012; Redondo *et al.*, 2014; Rosenbauer *et al.*, 2012; Scheuing *et al.*, 2015; Wang, Wiebe, & White, 2011; Willi *et al.*, 2015), puis par au moins cinq études depuis (Agarwal *et al.*, 2019; Bergenstal *et al.*, 2017; Mönkemöller *et al.*, 2019; Thomas *et al.*, 2018; Zuckerman-Levin *et al.*, 2018). L'ensemble de ces articles montre que les patients ayant une origine ethnique minoritaire ont un taux d'HbA1c plus élevé que les autres patients.

L'impact de l'origine des parents est particulièrement complexe à expliquer puisqu'il est corrélé à plusieurs autres facteurs comme le statut socio-économique (Chalew *et al.*, 2018; Mönkemöller *et al.*, 2019; Zuckerman-Levin *et al.*, 2018), la compréhension et l'adhérence des parents (Chalew *et al.*, 2018), l'utilisation d'un dispositif de suivi du diabète (Chalew *et al.*, 2018) et de facteurs biologiques (Bergenstal *et al.*, 2017; Chalew & Gomez, 2019). Dans notre



cohorte, seule la corrélation et l'interaction entre l'origine ethnique des patients et le statut socio-économique auraient pu être testée puisque nous n'avons pas collecté d'information sur les autres facteurs évoqués. Cependant, nous ne possédons ces deux informations que pour 106 des patients de la cohorte. Nous avons donc décidé de ne pas le tester pour le moment, même si l'impact du statut socio-professionnel de la mère semble être suggéré pour le taux d'HbA1c dans notre cohorte.

En revanche, les deux facteurs identifiés uniquement par l'étude de l'IDAA1C sont deux facteurs pour lesquels l'impact sur le taux d'HbA1c dans notre revue semble assez complexe puisque dans les deux cas, l'ensemble des études incluent dans notre revue de la littérature ne montre pas un effet ou une absence d'effet du type de traitement ou de la puberté (Gloaguen *et al.*, 2018), pouvant suggérer des effets assez faibles corrélés avec la dose d'insuline. Il serait donc intéressant de regarder les effets des interactions entre la DIR et le type de traitement, ainsi que la DIR et l'âge dans l'étude finale.

Il est important de noter que cette étude est une étude préliminaire sur les données disponibles et qu'il reste une partie des données à saisir, ce qui fait que nous manquons de puissance pour détecter certains facteurs, notamment pour la partie génétique. Pour identifier des facteurs génétiques à effets faibles, il est important de travailler sur de grands échantillons. Cela est possible grâce à l'étude de cohortes de patients comme le *Diabetes-patient-documentation* (DPV) en Allemagne et Autriche (Gerstl *et al.*, 2008) ou le *T1DX Exchange* aux Etats-Unis (Beck *et al.*, 2012), qui sont composées de plus de 10 000 patients. En France, aucune étude équivalente n'a été mise en place, mais l'autorisation depuis 2013 de l'utilisation des bases de données de santé nationale, comme le système national d'information inter-régimes de l'Assurance Maladie (SNIIRAM), pourrait permettre de créer ce type de très grande base de données pour le DT1.

De nouveaux outils ont depuis une dizaine d'années vu le jour, permettant pour les patients de mieux adapter leur traitement, ces évolutions permettent à la fois d'améliorer leur équilibre glycémique, diminuant ainsi le risque de complication, et leur qualité de vie (Zimmerman, Albanese-O'Neill, & Haller, 2019). Ces outils, et notamment les systèmes de surveillance continue de la glycémie, semblent associés à une diminution du taux d'HbA1c, aussi bien dans

le DPV en Allemagne que le T1DX aux Etats-Unis (DeSalvo *et al.*, 2018). De plus, parmi ces outils, plusieurs permettent de partager les données de glycémie ou d'insuline, notamment avec le clinicien. Ces outils pourraient permettre d'améliorer la collecte des données dans le cadre de projet de recherche, permettant d'étudier ces différents facteurs plus facilement et surtout de façon plus précise (Pralhad, Tanenbaum, Hood, & Maahs, 2018).

Une autre limite de notre étude est liée au fait qu'il s'agisse d'une étude rétrospective. La collecte des données peut s'avérer complexe dans ce type d'étude, notamment pour les données socio-économiques, de qualité de vie ou l'adhérence des patients au traitement qui sont pourtant des facteurs qui peuvent fortement influencer le taux d'HbA1c (Gloaguen *et al.*, 2018). De plus, les études rétrospectives entraînent la présence d'un certain nombre de perdus de vue et de données manquantes, limitant le nombre de sujets analysés.

Malgré tout, cette étude préliminaire a permis de répliquer l'effet de plusieurs facteurs déjà connus pour influencer le taux d'HbA1c chez l'enfant DT1 dans une cohorte de petite taille française, que sont la dose d'insuline requise et l'origine des parents. Une fois la collecte des données terminées, il sera important de refaire toutes ces analyses, mais aussi d'étudier plus en détails les profils d'évolution du taux d'HbA1c. En effet, pour le moment, nous avons surtout regardé quels étaient les profils que nous pouvions observer dans notre cohorte, sans identifier les facteurs qui pourraient être associés à ces trajectoires.

Nous avons ainsi pu montrer dans notre cohorte, qu'il ne semblait pas y avoir de lien entre le taux d'HbA1c au diagnostic et celui à 2 ans. Le côté prédictif du taux d'HbA1c au diagnostic ne semble pas clair ; en effet, dans notre revue, nous avons reporté que parmi les trois études qui avaient étudié l'effet du taux d'HbA1c au diagnostic, deux d'entre elles montraient un effet significatif, mais relativement faible (Chemtob, Hochhauser, Shemesh, Schmeidler, & Rapaport, 2011; Hanberger, Akesson, & Samuelsson, 2014), et une étude ne montrait aucune association entre le taux d'HbA1c au diagnostic et celui à un an (Redondo *et al.*, 2014). Depuis, une étude suédoise a montré une association entre le taux d'HbA1c durant le suivi du DT1 et celui au diagnostic, mais que celui-ci ne pouvait pas être utilisé comme prédicteur de l'équilibre glycémique au niveau individuel (Nilsson, Åkesson, Hanberger, & Samuelsson, 2017). À l'inverse, une étude publiée récemment a montré que le taux d'HbA1c moyen durant les trois

premiers mois de DT1 est un bon prédicteur de l'équilibre glycémique pour les 24 premiers mois de DT1, en utilisant des modèles polynomiaux fractionnés à effets mixtes (Mazarello Paes *et al.*, 2020). Le taux d'HbA1c au diagnostic n'aurait donc pas un effet linéaire sur le taux d'HbA1c sur le long terme, ce qu'il pourrait être intéressant de tester dans la cohorte GENDIR.

Pour le moment, aucune analyse n'a été effectuée sur les trajectoires du taux d'HbA1c que nous avons identifiées, puisque toutes les données à 5 et 10 ans n'ont pas été collectées. Plusieurs études récentes se sont intéressées aux facteurs pouvant influencer les trajectoires du taux d'HbA1c sur plusieurs années. Parmi ces études, deux d'entre elles ont effectuées des études en stratifiant leur échantillon en fonction de différentes caractéristiques. Les caractéristiques au diagnostic, telles que l'âge ou la présence d'une acidocétose, ne sembleraient pas influencer le taux d'HbA1c (Aeppli, Mahler, & Konrad, 2019). L'évolution du taux d'HbA1c semble aussi dépendre de l'ethnie du patient, avec une évolution plus défavorable dans les ethnies minoritaires (Redondo *et al.*, 2018). Cette stratégie pourrait être adaptée à nos données. Une seconde stratégie a été de modéliser l'ensemble des données collectées afin de créer des trajectoires, puis d'identifier par la suite les facteurs susceptibles d'expliquer l'appartenance d'un patient à une trajectoire particulière (Helgeson *et al.*, 2018). Cette étude a permis de montrer la présence de cinq trajectoires particulières, trois profils stables à des taux différents d'HbA1c, ainsi que deux profils d'augmentation du taux d'HbA1c pendant l'adolescence ou après 20 ans. Ces différents profils semblent être influencés par plusieurs facteurs psychologiques de l'enfants et des parents. Il semble en revanche compliqué de mettre en place ce type de stratégie dans les données GENDIR puisque nous n'avons collecté les données qu'à trois temps fixes. Il aurait fallu collecter un plus grand nombre de données afin de pouvoir faire une meilleure modélisation.

À terme, cette étude pourrait permettre de mettre en évidence des facteurs génétiques et non-génétiques qui permettront aux cliniciens d'avoir une meilleure compréhension des profils d'évolution de l'équilibre glycémique.

## Chapitre 5 | Discussion générale

---

L'objectif de ma thèse était d'identifier les facteurs influençant l'équilibre glycémique chez le jeune patient DT1 après deux ans de diabète, grâce à trois volets différents, qui peuvent être regroupés en deux axes :

- Un axe méthodologique concernant l'utilisation du GRM dans le cas d'un trait quantitatif ;
- L'étude de l'équilibre glycémique chez le jeune patient souffrant de DT1, qui est elle-même composée de deux parties : une revue de la littérature puis une étude des facteurs identifiés sur les données de GENDIR.

Chacun de ces deux axes seront discutés séparément, en mettant en avant les points positifs, ainsi que les réserves que nous pouvons avoir. Les perspectives de chacun de ces deux axes seront données en même temps.

### 5.1 Le GRM

Le premier volet de cette thèse consistait à étudier dans le cadre d'un trait quantitatif les propriétés statistiques du GRM, un test d'association génétique qui permet de tester l'association sans faire d'hypothèse sur le modèle génétique sous-jacent puis de le tester dans un second temps. Nous avons effectué une étude de simulation, qui a permis de montrer, d'une part, que ce test peut être plus puissant que le test d'association LR\_ADD classiquement utilisé qui fait l'hypothèse d'un modèle additif, et, d'autre part, que le modèle simulé est dans la plupart des cas le modèle retenu. Nous avons aussi comparé le GRM à deux tests récents, permettant de ne pas faire d'hypothèse sur le mode de transmission : un test allélique développé pour les traits quantitatifs (Lee *et al.*, 2013) et un test MAX basé sur une régression linéaire (Loley *et al.*, 2013). Dans les deux cas, le GRM était globalement plus puissant et plus rapide.

Nous avons ensuite appliqué ce test pour la recherche des facteurs génétiques influençant le taux d'HbA1c chez des individus non-diabétiques de la cohorte D.E.S.I.R.. La réanalyse de ces données a permis de suggérer un nouveau *locus* qui montre un écart à l'additivité. L'application

du GRM sur les données de la cohorte D.E.S.I.R. nous a ainsi permis d'illustrer le gain de puissance du GRM sur des données réelles pour détecter l'association par rapport au test LR\_ADD.

L'utilisation du modèle de régression linéaire sous modèle additif est de plus en plus controversée. En effet, aussi bien les études de simulations pour un trait quantitatif ou binaire (Dizier *et al.*, 2017; Gloaguen *et al.*, 2019), que pour des études sur des données de traits chez les plantes ou les animaux (Bonnafeous *et al.*, 2018; He *et al.*, 2015), ont montré une perte de puissance quand seul ce mode de transmission est testé. Nos résultats vont dans le même sens qu'une étude montrant que même si la majorité des marqueurs influençant les concentrations de métabolites chez l'Homme peut être détectée en utilisant la régression linéaire sous un modèle additif, l'utilisation d'un modèle génotypique, proche du GRM permet de détecter l'ensemble des marqueurs transmis de manière non-additive (Tsepilov *et al.*, 2015).

Une autre stratégie a été de développer des tests qui ne font pas *a priori* sur le mode de transmission, et notamment les tests MAX. En plus, du test MAX développé par Loley *et al.* (Loley *et al.*, 2013) et du test allélique pour les traits quantitatifs (Lee *et al.*, 2013), que nous avons comparé au GRM, différents tests ont été développés pour tester l'association sans *a priori* sur le mode de transmission du trait, dont un autre test MAX, basé sur les modèles marginaux (Kitsche, Ritz, & Hothorn, 2016), et un test basé sur un modèle bayésien de régression polynomiale (Bae, Perls, Steinberg, & Sebastiani, 2015). Aucun de ces deux nouveaux tests n'ont été comparés au GRM. Il pourrait être intéressant de le faire, pour voir l'ajout de ces deux méthodes innovantes, basées sur des concepts mathématiques plus complexes. D'autres méthodes MAX, basées sur des tests non paramétriques, ont aussi été développées (Konietschke, Libiger, & Hothorn, 2012; Li, Li, Zheng, Gao, & Yu, 2013; Zhang & Li, 2015), permettant ainsi de tester l'association dans le cas de traits non normalement distribués. Ces tests peuvent être particulièrement utiles pour tester l'association pour l'étude de traits non-normalisables. Cependant l'ensemble de ces tests semble avoir un temps de calcul très important, qui les rend difficilement utilisables dans le cadre de GWAS.

Les tests génotypiques sont peu utilisés, alors qu'ils sont implémentés dans la majorité des logiciels permettant d'effectuer un GWAS. Cela peut s'expliquer par leur plus grande complexité

d'interprétation, avec deux paramètres au lieu d'un seul, notamment en ce qui concerne le mode de transmission ou la taille de l'effet sur le trait étudié (Loley *et al.*, 2013).

L'autre raison expliquant la faible utilisation des tests génotypiques est leur perte de puissance dans le cas de marqueurs ayant un effet additif sur le phénotype ou la maladie étudiée, dû à l'ajout d'un degré de liberté. Or, comme il l'est indiqué dans le tutoriel sur les méthodes statistiques à adopter dans le cadre des études d'association pangénomiques (Balding, 2006), les premières études ont choisi la méthode qui semblait avoir la meilleure puissance. En absence de connaissance sur les modes de transmission des traits étudiés, la modélisation additive semblait permettre de détecter avec la meilleure puissance les marqueurs suivant un mode additif, qui semblent plus nombreux (Tsepilov *et al.*, 2015), et avec une puissance acceptable les marqueurs ayant un effet fort sous un modèle non-additif. Cependant, même si le GRM peut être moins puissant que la régression linéaire sous un modèle additif, cette perte de puissance est inférieure à 9%, lorsque que les gains de puissance maximum pour les modes de transmission dominant et récessif sont respectivement supérieurs à 24% et 80%. L'ajout d'un degré de liberté dû à l'incorporation d'une seconde variable dans le modèle est donc peu associé à une perte de puissance lorsque l'ensemble des modes de transmission est pris en compte.

Il pourrait être intéressant de développer le GRM pour d'autres types d'analyses, comme de la survie, comme cela a été proposé pour les tests MAX développés par Loley *et al* et Kitsche *et al* (Kitsche *et al.*, 2016; Loley *et al.*, 2013).

Un point crucial pour les GWAS actuellement est l'utilisation de cohorte de plus en plus grande pour pouvoir détecter des facteurs ayant des effets faibles (Sazonovs & Barrett, 2018). Afin d'atteindre la taille d'échantillon nécessaire (d'un peu moins de 10 000 individus par exemple d'après la revue de Sazonovs et Barrett pour un odd-ratio de 2 (Sazonovs & Barrett, 2018)), une des solutions consiste à effectuer des méta-analyses. Actuellement, la majorité de ces analyses est faite à l'aide du logiciel METAL (Willer, Li, & Abecasis, 2010), qui permet d'effectuer une analyse à partir des p-valeurs et de l'effet calculé dans l'étude de chacune des cohortes prises en compte dans la méta-analyse. Il pourrait donc être intéressant d'étendre METAL afin de pouvoir prendre en compte dans les méta-analyses les études effectuées avec le GRM.

Une autre solution utilisée pour augmenter le nombre de patients dans les études consiste à utiliser des registres ou des cohortes de patients, regroupant un très grand nombre d'informations afin de sélectionner des patients, qui donnent ensuite leur accord pour être génotypés. Cette stratégie a notamment été utilisée pour les études les plus récentes sur le taux d'HbA1c, permettant ainsi de collecter des données glycémiques sur plus de 10 000 participants, malades ou non (Kanai *et al.*, 2018; Spracklen *et al.*, 2018). En France, cette volonté d'accroître la recherche en génétique pour aider la médecine est particulièrement visible avec la mise en place depuis 2016 du Plan France Médecine Génomique 2025. A terme, ce plan devrait permettre d'intégrer le séquençage du génome en routine et dans la pratique clinique et de développer la recherche clinique dans ce domaine.

L'étude de plus en plus fréquente de variants rares s'explique notamment par l'augmentation du nombre d'études utilisant le séquençage de nouvelle génération (NGS pour *Next-Generation Sequencing*) (Wang, Lu, & Zhao, 2015). Différentes méthodes pour analyser les variants rares dans le cadre de GWAS sans passer par l'augmentation de la taille de l'échantillon ont été développées et ont fait l'objet de plusieurs revues (Lee, Abecasis, Boehnke, & Lin, 2014; Nicolae, 2016). La majorité de ces méthodes consiste à regrouper et à résumer l'information de plusieurs marqueurs pour une région donnée (un gène ou un bloc de LD) pour ensuite tester l'association à ce niveau. D'autres tests basés sur le lasso (pour *Least Absolute Shrinkage and Selection Operator* en anglais) et/ou sur les modèles bayésiens ont aussi été adaptés pour tester l'association entre des variants rares et des phénotypes. La méthode lasso étant basée sur un modèle de régression linéaire ou logistique en fonction du trait étudié, il serait tout à fait envisageable de l'adapter pour pouvoir modéliser en même temps le mode de transmission du marqueur afin de se rapprocher de l'idée du GRM.

Un autre développement intéressant pour le GRM serait l'intégration des méthodes d'analyse des interactions gène-gène ainsi que gène-environnement. En effet, de plus en plus d'études s'intéressent à l'intégration de ces types d'interactions dans les analyses (Cordell, 2009; DeWan, 2018; Gusareva & Van Steen, 2014; Ritz *et al.*, 2017). La méthode classique d'étude des interactions est basée sur l'ajout d'un terme d'interaction entre une covariable et le SNP étudié, PLINK a par ailleurs intégré cette méthode d'analyse notamment dans le cadre de l'analyse génotypique (Chang *et al.*, 2015; Purcell *et al.*, 2007). Cependant cette méthode, assez simple

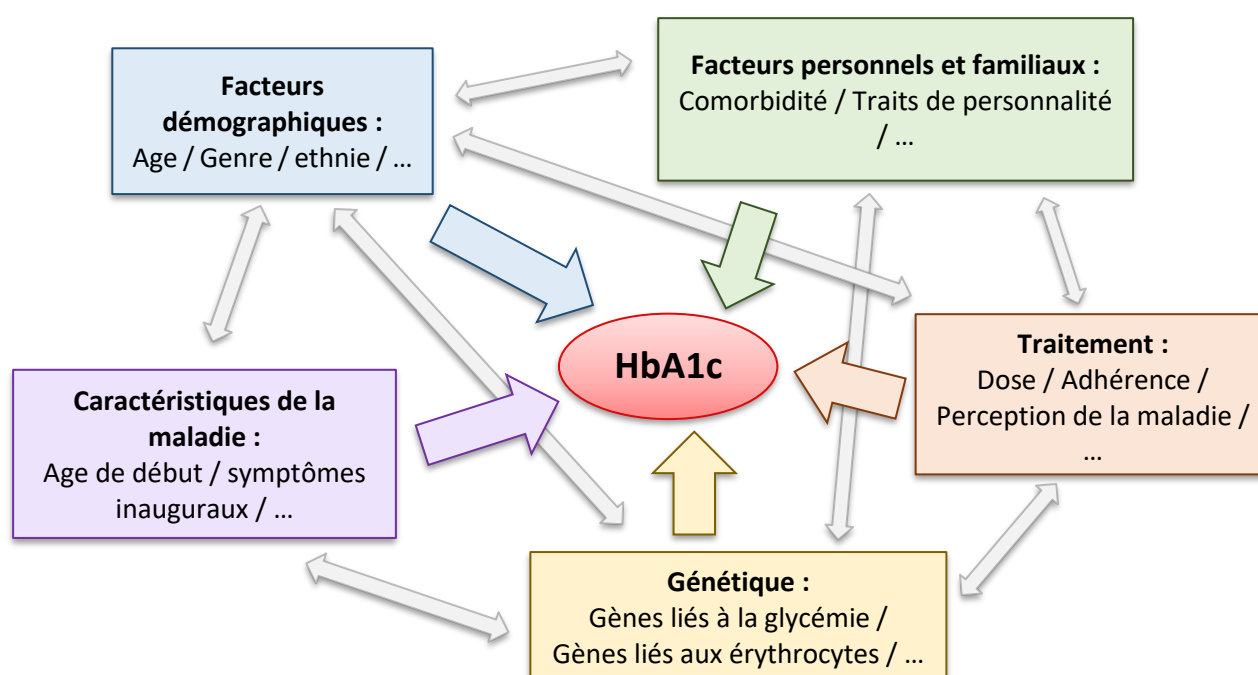


à mettre en place, nécessite des cohortes d'études de taille importante afin d'avoir une puissance suffisante pour détecter l'association des marqueurs et leur interaction avec l'environnement étudié (Simon, Sylvestre, Tremblay, & Hamet, 2016). Des méthodes plus complexes ont ainsi été proposées, afin de filtrer le nombre de marqueurs pour lesquels l'interaction sera testée, de réorganiser les données afin d'effectuer un plus petit nombre de tests, ou réduire les hypothèses des régressions linéaires, avec de nouveau l'utilisation du lasso (Gao *et al.*, 2013). Tout comme pour l'étude de variants rares, il serait envisageable d'adapter les méthodes lasso développées pour l'études des interactions gène-gène et gène-environnement afin de modéliser les différents modes de transmission génétique, comme le GRM. L'intégration de ces différentes interactions et l'augmentation du nombre de données non-génétiques, notamment par l'utilisation d'objets connectés, pourrait permettre sur le long terme une médecine personnalisée de précision, mais nécessitera une évolution importante des méthodes d'analyse, notamment en utilisant des méthodes d'intelligence artificielle et l'inférence bayésienne (Li, Li, Zhang, & Snyder, 2019).

## 5.2 Le taux d'HbA1c chez le jeune patient souffrant de DT1

Afin d'étudier l'équilibre glycémique du jeune patient souffrant de DT1, nous avons effectué deux études distinctes, composant les deuxième et troisième volets de ma thèse.

Le deuxième volet visait à identifier les facteurs connus pour influencer le taux d'HbA1c chez les enfants, adolescents et jeunes adultes atteints de DT1 depuis au moins deux ans et de synthétiser cela dans une revue systématique de la littérature. Cette revue a permis de montrer la complexité du taux d'HbA1c chez le jeune patient DT1. Comme montré dans le modèle simplifié (Figure 26), un certain nombre d'entre eux sont corrélés entre eux.



**Figure 26 :** Modèle simplifié des différents facteurs impliqués dans la régulation du taux d'HbA1c chez les patients souffrant de DT1

Les flèches grises représentent les corrélations qui peuvent exister entre les différents facteurs influençant le taux d'HbA1c, adapté de Florez (Florez, 2010) et de notre revue de la littérature (Gloaguen, Bendelac, Nicolino, Julier, & Mathieu, 2018)

Le troisième volet consistait à participer à une étude, qui a comme objectif de tester ces différents facteurs non-génétiques et génétiques dans une cohorte de patients français souffrant de DT1. L'étude préliminaire de cette cohorte a permis de mettre en évidence l'impact de l'origine géographique des parents.

Ces deux volets de ma thèse ont ainsi permis de mettre en évidence la complexité du taux d'HbA1c chez l'enfant et le jeune adulte atteint de DT1. Le taux d'HbA1c est influencé par un

grand nombre de facteurs non-génétiques, qui sont en partie corrélés entre eux, faisant du taux d'HbA1c un facteur complexe à étudier. Le taux d'HbA1c reste toutefois un marqueur clé pour le diagnostic et le suivi du diabète, notamment comme marqueur du développement de complications (ADA, 2019a).

Prendre en compte les facteurs influençant le taux d'HbA1c pourraient permettre de mieux expliquer pourquoi certains patients n'arrivent jamais à être bien équilibrés malgré un traitement bien conduit, et pourquoi à l'inverse, d'autres, malgré une mauvaise adhérence au traitement, arrivent à avoir un taux d'HbA1c stable et inférieur à 7.5%. Une meilleure identification des caractéristiques de ces patients pourrait permettre de mettre en place des traitements spécifiques. En effet, certaines formes de diabète atypiques, d'origine monogénique, peuvent avoir des traitements différents (Kleinberger & Pollin, 2015). Il pourrait être envisageable que certains patients ayant des caractéristiques bien particulières puissent bénéficier de traitements personnalisés.

Une seule étude s'est pour le moment intéressée à l'impact des facteurs identifiés sur la variance du taux d'HbA1c. Cette étude se déroulait après la participation à un programme intensif sur le traitement du diabète. Cette étude a montré que le fait d'être une femme, de fumer, d'avoir été diagnostiqué DT1 jeune, de faire peu de mesure de la glycémie chez soi, d'avoir un statut socio-économique faible, d'avoir de faible connaissance sur le DT1, et de moins supporter la maladie, expliquent seulement 17% de la variance du taux d'HbA1c (Bott *et al.*, 1994). Cependant, cette étude est ancienne, même si elle permet de montrer que les facteurs identifiés aujourd'hui expliquent peu de la variance du taux d'HbA1c.

Le principal facteur identifié aussi bien dans la revue que dans l'étude de GENDIR est l'ethnie, qui semblent avoir un rôle majeur dans le contrôle du taux d'HbA1c. Une revue récente a montré que le taux d'HbA1c chez des individus non diabétiques d'origine caucasienne est en moyenne plus bas que le taux d'HbA1c chez des individus d'origines asiatique, africaine et latino-américaine, ce qui rend le seuil de diagnostic de diabète de 6.5% dépendant de l'origine du patient (Cavagnoli, Pimentel, Freitas, Gross, & Camargo, 2017). Il est donc important de développer et d'adapter les seuils en fonction des caractéristiques des patients. Choisir un seuil adapté aussi bien pour le diagnostic du diabète que pour le suivi semble donc un point crucial pour la suite des études sur le taux d'HbA1c.

Une des explications possibles de l'impact de l'ethnie sur le taux d'HbA1c est la présence de maladies connues pour influencer celui-ci, notamment dans la population africaine, telle que la drépanocytose, les néphropathies chroniques et les infections par le VIH, rendant l'utilisation du taux d'HbA1c complexe pour le diagnostic et le suivi du diabète (George & Erasmus, 2018; Krhac & Lovrencic, 2019). Pour répondre à ce problème, d'autres biomarqueurs de la glycémie, tels que l'albumine glyquée et la fructosamine, ont été proposés (George & Erasmus, 2018; Krhac & Lovrencic, 2019). Leur utilisation pourrait être aussi intéressante chez les femmes enceintes, qui ont une variation de leur taux d'HbA1c (Krhac & Lovrencic, 2019). Cependant, tout comme le taux d'HbA1c, ces biomarqueurs sont influencés par des facteurs génétiques glycémiques et non-glycémiques, qui peuvent aussi être dépendants de l'origine ethnique des patients (Loomis *et al.*, 2018). De plus, les méthodes de mesures de ces biomarqueurs ne sont pas encore utilisées en routine et peuvent être plus coûteuses (Krhac & Lovrencic, 2019). L'utilisation conjointe de l'albumine glyquée et du taux d'HbA1c, permettraient de mieux suivre l'évolution du DT1, notamment dans les populations asiatiques, permettant ainsi un meilleur management du DT1 (Musha *et al.*, 2018).

Notre étude a porté sur des enfants et jeunes adultes, mais d'autres études sur d'autres périodes de la vie ont identifié d'autres facteurs, ou les mêmes facteurs avec des effets différents. Parmi les facteurs déjà identifiés chez le patient jeune, l'effet de deux facteurs semblent particulièrement différent. Premièrement, l'effet de l'âge sur le taux d'HbA1c n'est pas linéaire, puisqu'il est associé à une augmentation du taux d'HbA1c jusqu'à 18 ans, puis une diminution jusqu'à 30 ans, et une stabilisation jusqu'à la fin de vie du patient (Miller *et al.*, 2015). L'utilisation de la pompe semble aussi avoir un impact différent chez le jeune par rapport à son utilisation chez l'adulte : alors que chez l'enfant, son utilisation est associée à un taux d'HbA1c plus faible, chez le patient adulte, le taux d'HbA1c est plus élevé chez les patients utilisant une pompe (Willers *et al.*, 2018). D'autres facteurs non recherchés chez l'enfant ont aussi un impact sur le taux d'HbA1c, tel que le statut fumeur (Gerber, Locher, Schmid, Spinass, & Lehmann, 2013).

Il pourrait donc être intéressant d'effectuer des analyses sur l'ensemble des âges de la vie pour identifier les facteurs ayant un impact constant, et dans un second d'effectuer des analyses en stratifiant sur l'âge, en prenant notamment les différentes tranches d'âge mises en évidence

par Miller *et al* (Miller *et al.*, 2020; Miller *et al.*, 2015 ). Par ailleurs, le nombre d'études s'intéressant aux facteurs influençant le taux d'HbA1c sur l'ensemble de la vie du patients ayant un DT1 a particulièrement augmenté depuis la publication de notre revue (Agarwal *et al.*, 2019; Clements *et al.*, 2019; Fortins *et al.*, 2019; Hermann *et al.*, 2019; Mair *et al.*, 2019; Nielsen, Gaulke, Eriksen, Svensson, & Skipper, 2019; Schwandt *et al.*, 2017; Teló, Dougher, Volkening, Katz, & Laffel, 2018; van Esdonk, Tai, Cotterill, Charles, & Hennig, 2017).

Concernant les facteurs génétiques, à ce jour, seuls les 55 marqueurs identifiés précédemment comme étant associés au taux d'HbA1c chez le sujet non-diabétique dans la méta-analyse de Wheeler *et al* ont été testés dans le contexte du taux d'HbA1c chez le patients DT1 (Syreeni *et al.*, 2019; Wheeler *et al.*, 2017). Seulement trois marqueurs, présents dans les gènes HK1, HFE et FOYN2 classés tous trois dans la catégorie érythrocytaire, ont été répliqués dans leur cohorte de patient DT1 (Syreeni *et al.*, 2019). Aucun des marqueurs identifiés chez les patients DT1 n'ont en revanche été spécifiquement testés, et n'ont pas non plus été identifiés lors des différents GWAS effectués sur les individus non-diabétiques. Nous avons donc d'un côté, chez les patients non-diabétiques, deux voies qui semblent principalement impliqué dans le contrôle du taux d'HbA1c : une voie érythrocytaire et une voie glycémique ; et de l'autre, chez les patients atteints de DT1, les gènes plutôt impliqués dans les voies immunologiques.

Le taux d'HbA1c permet de refléter la glycémie sur trois mois, mais peut aussi être en partie un indicateur de l'activité résiduelle des cellules  $\beta$  des ilots de Langerhans du pancréas, comme le montre l'IDAA1C (Mortensen *et al.*, 2009). Chez les patients DT1, cette destruction des cellules  $\beta$  est dû à une réaction auto-immune. Il serait donc envisageable que la présence de plusieurs gènes impliqués dans le système immunitaire parmi les gènes associés au taux d'HbA1c chez les patients DT1 puisse s'expliquer par le fait que ces gènes sont aussi impliqués dans la réaction auto-immune à l'origine du DT1.

Les marqueurs génétiques identifiés à ce jour n'expliquent qu'une faible part de la part génétique du taux d'HbA1c, aussi bien les facteurs identifiés chez les patients non-diabétiques que ceux souffrant de diabète : Wheeler *et al* rapportent que ces 60 régions identifiées chez les non-diabétiques expliquent entre 4 et 14% de la variance du taux d'HbA1c (Wheeler *et al.*, 2017) ; et chez les patients souffrant de diabète, les différentes régions identifiées par Paterson *et al* expliquent moins de 2% de la variance totale (Paterson *et al.*, 2010). Le peu de variance

expliquée par ces marqueurs pose la question de l'importance de ces variants dans le cadre de la médecine personnalisée. En effet, même si d'un point de vue recherche, il est toujours intéressant d'identifier les différents mécanismes qui sont impliqués dans une maladie, d'un point de vue médical, il peut être difficile d'intégrer ces résultats dans le parcours de soin lorsque les effets sont faibles.

Il existe certaines formes de diabète, dites monogéniques, qui peuvent mimer un DT1, mais qui sont dues en réalité à une ou plusieurs mutations dans un unique gène. Les patients souffrant de certaines formes monogéniques peuvent ne pas avoir besoin d'insuline pour équilibrer leur diabète. Les principales formes monogéniques de diabète et leur prise en charge ont été décrites dans une revue publiée en 2018 (Sanyoura, Philipson, & Naylor, 2018). Par exemple, les diabètes causés par une mutation dans les gènes *KCNJ11* et *ABCC8*, codant pour les sous-unités des canaux potassiques impliqués dans la sécrétion de l'insuline, peuvent être traités par sulfamides hypoglycémisants, traitement utilisé pour le DT2 (Klee *et al.*, 2012). Il semble donc intéressant de poursuivre des études à grande échelle de type GWAS pour identifier les gènes à effet faible, et en parallèle d'étudier les variants rares, notamment dans des familles consanguines et multiplexes (Zalloua *et al.*, 2008). C'est l'objet d'ailleurs d'un des 4 projets pilotes du Plan France Médecine Génomique 2025.

Un autre point important est le développement de capteur permettant de mesurer de façon continue la glycémie des patients, mais aussi le développement d'applications mobiles permettant une lecture simple et à tout moment de la glycémie. Ces nouvelles technologies pourraient permettre à certains patients de mieux réguler leur glycémie. Pour le moment, l'impact de ces technologies reste faible ou non mesurable sur le taux d'HbA1c (Knox, Quirk, Glazebrook, Randell, & Blake, 2019). Ces nouveaux objets pourraient aussi permettre d'étudier plus finement l'effet de la dose d'insuline sur le taux d'HbA1c. Aujourd'hui, dans les études, les données collectées correspondent aux doses d'insuline prescrites par le praticien, alors qu'il serait plus intéressant de regarder l'impact des doses réellement prises par le patient.

Malgré tout, un point critique de notre stratégie est le fait que nous avons effectué une étude rétrospective. En effet, cette stratégie possède certains points positifs, notamment sur le fait de pouvoir avoir accès à des données longitudinales, puisque nous avons collecté des données à 10 ans de diabète. Cependant, le travail de collecte des données est couteux en temps en

raison de l'absence de base de données structurées. En effet, les données collectées n'étaient pas référencées dans la base de données informatisée, ce qui nous a obligé de recueillir les données dans les dossiers papiers aux archives des Hospices Civils de Lyon.

Cette stratégie pourrait être complétée par l'utilisation des bases de données de santé nationale, comme l'utilisation du SNIIRAM ou des données de cohortes de patients. L'utilisation de cohorte prospective, avec notamment l'étude Constance en France, qui suit en France 200 000 adultes issus de la population générale (Zins & Goldberg, 2015) pourrait permettre d'identifier les facteurs impliqués dans le contrôle du taux d'HbA1c.

Enfin, il serait également intéressant d'identifier et de tester l'impact de facteurs non-génétiques et des marqueurs génétiques identifiés aussi bien chez les personnes non-diabétiques que diabétiques, en prenant en compte les interactions possibles entre les facteurs génétiques et environnementaux. L'interaction entre certains facteurs génétiques et non-génétiques pourraient permettre d'expliquer une plus grande part de la variance du taux d'HbA1c. Ces études seraient envisageables en utilisant les futurs développements du GRM.



# Bibliographie

---

- Abali, Turan, Atay, Guran, Haliloglu, & Bereket. (2015). Higher insulin detemir doses are required for the similar glycemic control: comparison of insulin detemir and glargine in children with type 1 diabetes mellitus. *Pediatr Diabetes, 16*(5), 361-366. doi:10.1111/pedi.12167
- Abdul-Rasoul, Habib, & Al-Khouly. (2006). 'The honeymoon phase' in children with type 1 diabetes mellitus: frequency, duration, and influential factors. *Pediatr Diabetes, 7*(2), 101-107. doi:10.1111/j.1399-543X.2006.00155.x
- Abecasis, Altshuler, Auton, Brooks, Durbin, Gibbs, *et al.* (2010). A map of human genome variation from population-scale sequencing. *Nature, 467*(7319), 1061-1073. doi:10.1038/nature09534
- Abu El-Ella, Khattab, El-Mekawy, & El-Shamy. (2018). CD226 gene polymorphism (rs763361 C>T) is associated with susceptibility to type 1 diabetes mellitus among Egyptian children. *Arch Pediatr, 25*(6), 378-382. doi:10.1016/j.arcped.2018.06.009
- ADA. (2011). Diagnosis and classification of diabetes mellitus. *Diabetes Care, 34 Suppl 1*, S62-69. doi:10.2337/dc11-S062
- ADA. (2016). Standards of Medical Care in Diabetes-2016: Summary of Revisions. *Diabetes Care, 39 Suppl 1*, S4-5. doi:10.2337/dc16-S003
- ADA. (2017). 2. Classification and Diagnosis of Diabetes. *Diabetes Care, 40*(Supplement 1), S11-S24. doi:10.2337/dc17-S005
- ADA. (2019a). 6. Glycemic Targets: Standards of Medical Care in Diabetes-2019. *Diabetes Care, 42*(Suppl 1), S61-s70. doi:10.2337/dc19-S006
- ADA. (2019b). Standards of Medical Care in Diabetes—2019 : 2. Classification and Diagnosis of Diabetes. *Diabetes Care, 42*(Supplement 1), S13-S28. doi:10.2337/dc19-S002
- Aeppli, Mahler, & Konrad. (2019). Future glycemic control of children diagnosed with type 1 diabetes mellitus at toddler and preschool/school age. *Journal of pediatric endocrinology & metabolism : JPEM, 32*(9), 929-933. doi:10.1515/jpem-2019-0152
- Agarwal, Khokhar, Castells, Marwa, Hagerty, Dunkley, *et al.* (2019). Role of Social Factors in Glycemic Control Among African American Children and Adolescents with Type 1 Diabetes. *Journal of the National Medical Association, 111*(1), 37-45. doi:10.1016/j.jnma.2018.05.002
- Akesson, Hanberger, & Samuelsson. (2015). The influence of age, gender, insulin dose, BMI, and blood pressure on metabolic control in young patients with type 1 diabetes. *Pediatr Diabetes, 16*(8), 581-586. doi:10.1111/pedi.12219
- Allen, A., & Balog. (1958). Observations on the Chromatographic Heterogeneity of Normal Adult and Fetal Human Hemoglobin: A Study of the Effects of Crystallization and Chromatography on the Heterogeneity and Isoleucine Content. *Journal of the American Chemical Society, 80*.
- Almasi, Aliparasti, Aliasgarzadeh, Abd-Nikfarjam, Hyoty, Knip, *et al.* (2015). Investigation of CTLA-4-318C/T gene polymorphism in cases with type 1 diabetes of Azerbaijan, Northwest Iran. *Immunology Letters, 166*(2), 134-139. doi:10.1016/j.imlet.2015.05.021
- Alsaleh, Smith, Keady, & Taylor. (2010). Insulin pumps: from inception to the present and toward the future. *Journal of Clinical Pharmacy and Therapeutics, 35*(2), 127-138. doi:10.1111/j.1365-2710.2009.01048.x

- Altshuler, Gibbs, Peltonen, Dermitzakis, Schaffner, Yu, *et al.* (2010). Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311), 52-58. doi:10.1038/nature09298
- An, Miljkovic, Thyagarajan, Kraja, Daw, Pankow, *et al.* (2014). Genome-wide association study identifies common loci influencing circulating glycosylated hemoglobin (HbA1c) levels in non-diabetic subjects: the Long Life Family Study (LLFS). *Metabolism*, 63(4), 461-468. doi:10.1016/j.metabol.2013.11.018
- Andreasson, Ekelund, Landin-Olsson, & Nilsson. (2018). HbA1c levels in children with type 1 diabetes and correlation to diabetic retinopathy. *Journal of pediatric endocrinology & metabolism : JPEM*, 31(4), 369-374. doi:10.1515/jpem-2017-0417
- Appel, Moltke, Jorgensen, Bjerregaard, Linneberg, Pedersen, *et al.* (2018). Genetic determinants of glycosylated hemoglobin levels in the Greenlandic Inuit population. *Eur J Hum Genet*. doi:10.1038/s41431-018-0109-3
- Assmann, Brondani Lde, Bauer, Canani, & Crispim. (2014). Polymorphisms in the TLR3 gene are associated with risk for type 1 diabetes mellitus. *Eur J Endocrinol*, 170(4), 519-527. doi:10.1530/EJE-13-0963 [doi]
- Auton, Brooks, Durbin, Garrison, Kang, Korbel, *et al.* (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68-74. doi:10.1038/nature15393
- Bachle, Lange, Stahl-Pehe, Castillo, Holl, Giani, *et al.* (2015). Associations between HbA1c and depressive symptoms in young adults with early-onset type 1 diabetes. *Psychoneuroendocrinology*, 55, 48-58. doi:10.1016/j.psyneuen.2015.01.026
- Bachle, Stahl-Pehe, & Rosenbauer. (2016). Disordered eating and insulin restriction in youths receiving intensified insulin treatment: Results from a nationwide population-based study. *Int J Eat Disord*, 49(2), 191-196. doi:10.1002/eat.22463
- Bae, Perls, Steinberg, & Sebastiani. (2015). Bayesian Polynomial Regression Models to Fit Multiple Genetic Models for Quantitative Traits. *Bayesian Anal*, 10(1), 53-74. doi:10.1214/14-BA880
- Bae, Suh, Jin, Kim, Hur, Kim, *et al.* (2014). Hemoglobin A1c values are affected by hemoglobin level and gender in non-anemic Koreans. *J Diabetes Investig*, 5(1), 60-65. doi:10.1111/jdi.12123
- Balding. (2006). A tutorial on statistical methods for population association studies. *Nat Rev Genet*, 7(10), 781-791. doi:10.1038/nrg1916
- Balkau. (1996). An epidemiologic survey from a network of French Health Examination Centres, (D.E.S.I.R.): epidemiologic data on the insulin resistance syndrome. *Rev Epidemiol Sante Publique*, 44(4), 373-375.
- Balkau, Soulimane, Lange, Gautier, Tichet, & Vol. (2011). Are the same clinical risk factors relevant for incident diabetes defined by treatment, fasting plasma glucose, and HbA1c? *Diabetes Care*, 34(4), 957-959. doi:10.2337/dc10-1581
- Barat. (2016). Epidemiology of type 1 diabetes in children. *Soins PEDIATR Pueric*(288), 10-12. doi:10.1016/j.spp.2015.11.003
- Barrett, Clayton, Concannon, Akolkar, Cooper, Erlich, *et al.* (2009). Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*, 41(6), 703-707. doi:10.1038/ng.381
- Baucom, Queen, Wiebe, Turner, Wolfe, Godbey, *et al.* (2015). Depressive symptoms, daily stress, and adherence in late adolescents with type 1 diabetes. *Health Psychol*, 34(5), 522-530. doi:10.1037/hea0000219

- Beck, Tamborlane, Bergenstal, Miller, DuBose, Hall, *et al.* (2012). The T1D Exchange clinic registry. *J Clin Endocrinol Metab*, 97(12), 4383-4389. doi:10.1210/jc.2012-1561
- Benjamini, & Hochberg. (1995). *Controlling The False Discovery Rate - A Practical And Powerful Approach To Multiple Testing* (Vol. 57).
- Bennett, Guo, & Dharmage. (2007). HbA(1c) as a screening tool for detection of Type 2 diabetes: a systematic review. *Diabet Med*, 24(4), 333-343. doi:10.1111/j.1464-5491.2007.02106.x
- Beraki, Magnuson, Sarnblad, Aman, & Samuelsson. (2014). Increase in physical activity is associated with lower HbA1c levels in children and adolescents with type 1 diabetes: results from a cross-sectional study based on the Swedish pediatric diabetes quality registry (SWEDIABKIDS). *Diabetes Res Clin Pract*, 105(1), 119-125. doi:10.1016/j.diabres.2014.01.029
- Bergenstal, Gal, Connor, Gubitosi-Klug, Kruger, Olson, *et al.* (2017). Racial Differences in the Relationship of Glucose Concentrations and Hemoglobin A1c Levels. *Annals of internal medicine*, 167(2), 95-102. doi:10.7326/M16-2596
- Birkebaek, Kahlert, Bjarnason, Drivvoll, Johansen, Konradsdottir, *et al.* (2018). Body mass index standard deviation score and obesity in children with type 1 diabetes in the Nordic countries. HbA(1c) and other predictors of increasing BMISDS. *Pediatr Diabetes*, 19(7), 1198-1205. doi:10.1111/pedi.12693
- Bishop, & Williamson. (1990). The power of identity-by-state methods for linkage analysis. *American journal of human genetics*, 46(2), 254-265.
- Bizzarri, Benevento, Patera, Bongiovanni, Boiani, Fusco, *et al.* (2013). Residual beta-cell mass influences growth of prepubertal children with type 1 diabetes. *Horm Res Paediatr*, 80(4), 287-292. doi:10.1159/000355116 [doi]
- Blackman, Raghinaru, Adi, Simmons, Ebner-Lyon, Chase, *et al.* (2014). Insulin pump use in young children in the T1D Exchange clinic registry is associated with lower hemoglobin A1c levels than injection therapy. *Pediatr Diabetes*, 15(8), 564-572. doi:10.1111/pedi.12121
- Blair, McKay, Ridyard, Thornborough, Bedson, Peak, *et al.* (2019). Continuous subcutaneous insulin infusion versus multiple daily injection regimens in children and young people at diagnosis of type 1 diabetes: pragmatic randomised controlled trial and economic evaluation. *BMJ (Clinical research ed.)*, 365, 11226-11226. doi:10.1136/bmj.11226
- Bode, Iotova, Kovarenko, Laffel, Rao, Deenadayalan, *et al.* (2019). Efficacy and Safety of Fast-Acting Insulin Aspart Compared With Insulin Aspart, Both in Combination With Insulin Degludec, in Children and Adolescents With Type 1 Diabetes: The onset 7 Trial. *Diabetes Care*, 42(7), 1255-1262. doi:10.2337/dc19-0009
- Bonafous, Fievet, Blanchet, Boniface, Carrère, Gouzy, *et al.* (2018). Comparison of GWAS models to identify non-additive genetic control of flowering time in sunflower hybrids. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 131(2), 319-332. doi:10.1007/s00122-017-3003-4
- Bonnefond, Clement, Fawcett, Yengo, Vaillant, Guillaume, *et al.* (2012). Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nat Genet*, 44(3), 297-301. doi:10.1038/ng.1053
- Bonnefond, Vaxillaire, Labrune, Lecoeur, Chevre, Bouatia-Naji, *et al.* (2009). Genetic variant in HK1 is associated with a proanemic state and A1C but not other glycemic control-related traits. *Diabetes*, 58(11), 2687-2697. doi:10.2337/db09-0652
- Bonney. (1984). On the statistical determination of major gene mechanisms in continuous human traits: regressive models. *Am J Med Genet*, 18(4), 731-749. doi:10.1002/ajmg.1320180420

- Bonney. (1986). A note on the basis of regressive models for genetic analysis. *Genet Epidemiol Suppl*, 1, 37-42.
- Bookchin, & Gallop. (1968). Structure of hemoglobin A1c: nature of the N-terminal beta chain blocking group. *Biochem Biophys Res Commun*, 32(1), 86-93. doi:0006-291X(68)90430-0 [pii]
- Boomsma, Busjahn, & Peltonen. (2002). Classical twin studies and beyond. *Nat Rev Genet*, 3(11), 872-882. doi:10.1038/nrg932
- Botros, Islam, & Hursh. (2019). Insulin pump therapy, pre-pump hemoglobin A(1c) and metabolic improvement in children with type 1 diabetes at a tertiary Canadian children's hospital. *Pediatr Diabetes*, 20(4), 427-433. doi:10.1111/pedi.12834
- Bott, Jörgens, Grüsser, Bender, Mühlhauser, & Berger. (1994). Predictors of glycaemic control in type 1 diabetic patients after participation in an intensified treatment and teaching programme. *Diabetic medicine : a journal of the British Diabetic Association*, 11(4), 362-371. doi:10.1111/j.1464-5491.1994.tb00287.x
- Bouatia-Naji, Bonnefond, Cavalcanti-Proenca, Sparso, Holmkvist, Marchand, *et al.* (2009). A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nat Genet*, 41(1), 89-94. doi:10.1038/ng.277
- Bouatia-Naji, Rocheleau, Van Lommel, Lemaire, Schuit, Cavalcanti-Proenca, *et al.* (2008). A polymorphism within the G6PC2 gene is associated with fasting plasma glucose levels. *Science*, 320(5879), 1085-1088. doi:10.1126/science.1156849
- Bouhanick, Gallois, Hadjadj, Boux de Casson, Limal, & Marre. (1999). Relationship between glomerular hyperfiltration and ACE insertion/deletion polymorphism in type 1 diabetic children and adolescents. *Diabetes Care*, 22(4), 618-622. doi:10.2337/diacare.22.4.618
- Bradfield, Qu, Wang, Zhang, Sleiman, Kim, *et al.* (2011). A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet*, 7(9), e1002293. doi:10.1371/journal.pgen.1002293
- Brancato, Fleres, Aiello, Saura, Scorsone, Ferrara, *et al.* (2014). The effectiveness and durability of an early insulin pump therapy in children and adolescents with type 1 diabetes mellitus. *Diabetes Technol Ther*, 16(11), 735-741. doi:10.1089/dia.2014.0034 [doi]
- Brorsson, Nielsen, Andersen, Kaur, Bergholdt, Hansen, *et al.* (2016). Genetic Risk Score Modelling for Disease Progression in New-Onset Type 1 Diabetes Patients: Increased Genetic Load of Islet-Expressed and Cytokine-Regulated Candidate Genes Predicts Poorer Glycemic Control. *Journal of diabetes research*. doi:10.1155/2016/9570424
- Brorsson, Viklund, Ortvist, & Lindholm Olinde. (2015). Does treatment with an insulin pump improve glycaemic control in children and adolescents with type 1 diabetes? A retrospective case-control study. *Pediatr Diabetes*, 16(7), 546-553. doi:10.1111/pedi.12209 [doi]
- Buchberger, Huppertz, Krabbe, Lux, Mattivi, & Siafarikas. (2016). Symptoms of depression and anxiety in youth with type 1 diabetes: A systematic review and meta-analysis. *Psychoneuroendocrinology*, 70, 70-84. doi:10.1016/j.psyneuen.2016.04.019
- Bunn, Haney, Gabbay, & Gallop. (1975). Further identification of the nature and linkage of the carbohydrate in hemoglobin A1c. *Biochemical and Biophysical Research Communications*, 67(1), 103-109. doi:10.1016/0006-291X(75)90289-2
- Bush, & Moore. (2012). Chapter 11: Genome-wide association studies. *PLoS Comput Biol*, 8(12), e1002822. doi:10.1371/journal.pcbi.1002822
- Butwicka, Zalepa, Fendler, Szadkowska, & Mlynarski. (2013). Maternal depressive symptoms predict acute hospitalization among children with type 1 diabetes. *Pediatr Diabetes*, 14(4), 288-294. doi:10.1111/pedi.12010 [doi]

- Cadario, Cerutti, Savastio, Rabbone, Tumini, & Bruno. (2014). Increasing burden, younger age at onset and worst metabolic control in migrant than in Italian children with type 1 diabetes: an emerging problem in pediatric clinics. *Acta Diabetol*, 51(2), 263-267. doi:10.1007/s00592-013-0514-6 [doi]
- Campbell, Schatz, Chen, Wong, Steck, Tamborlane, *et al.* (2014). A contrast between children and adolescents with excellent and poor control: the T1D Exchange clinic registry experience. *Pediatr Diabetes*, 15(2), 110-117. doi:10.1111/pedi.12067 [doi]
- Carlson, Eberle, Rieder, Yi, Kruglyak, & Nickerson. (2004). Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet*, 74(1), 106-120. doi:10.1086/381000
- Cavagnoli, Comerlato, Comerlato, Renz, Gross, & Camargo. (2011). HbA(1c) measurement for the diagnosis of diabetes: is it enough? *Diabet Med*, 28(1), 31-35.
- Cavagnoli, Pimentel, Freitas, Gross, & Camargo. (2015). Factors affecting A1C in non-diabetic individuals: Review and meta-analysis. *Clin Chim Acta*, 445, 107-114. doi:10.1016/j.cca.2015.03.024
- Cavagnoli, Pimentel, Freitas, Gross, & Camargo. (2017). Effect of ethnicity on HbA1c levels in individuals without diabetes: Systematic review and meta-analysis. *PLoS One*, 12(2), e0171315. doi:10.1371/journal.pone.0171315
- Cengiz, Xing, Wong, Wolfsdorf, Haymond, Rewers, *et al.* (2013). Severe hypoglycemia and diabetic ketoacidosis among youth with type 1 diabetes in the T1D Exchange clinic registry. *Pediatr Diabetes*, 14(6), 447-454. doi:10.1111/pedi.12030
- Chalew, & Gomez. (2019). A labile form of hemoglobin A1c is higher in African-American youth with type 1 diabetes compared to Caucasian patients at similar glucose levels. *Pediatr Diabetes*, 20(6), 736-742. doi:10.1111/pedi.12863
- Chalew, Gomez, Vargas, Kamps, Jurgen, Scribner, *et al.* (2018). Hemoglobin A1c, frequency of glucose testing and social disadvantage: Metrics of racial health disparity in youth with type 1 diabetes. *Journal of diabetes and its complications*, 32(12), 1085-1090. doi:10.1016/j.jdiacomp.2018.02.008
- Chalew, McCarter, & Hempe. (2013). Biological variation and hemoglobin A1c: relevance to diabetes management and complications. *Pediatr Diabetes*, 14(6), 391-398. doi:10.1111/pedi.12055
- Chang, Chow, Tellier, Vattikuti, Purcell, & Lee. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4, 7. doi:10.1186/s13742-015-0047-8
- Chemtob, Hochhauser, Shemesh, Schmeidler, & Rapaport. (2011). Does poor early metabolic control predict subsequent poor control in young children with type 1 diabetes: an exploratory study. *Journal of Diabetes*, 3(2), 153-157. doi:10.1111/j.1753-0407.2011.00116.x
- Chen, Ong, Tay, Sim, Ali, Xu, *et al.* (2013). A study assessing the association of glycosylated hemoglobin A1C (HbA1C) associated variants with HbA1C, chronic kidney disease and diabetic retinopathy in populations of Asian ancestry. *PLoS One*, 8(11), e79767. doi:10.1371/journal.pone.0079767
- Chen, Takeuchi, Lee, Li, Wu, Liang, *et al.* (2014). Multiple nonglycemic genomic loci are newly associated with blood level of glycosylated hemoglobin in East Asians. *Diabetes*, 63(7), 2551-2562. doi:10.2337/db13-1815
- Cherubini, Pintaudi, Rossi, Lucisano, Pellegrini, Chiumello, *et al.* (2014). Severe hypoglycemia and ketoacidosis over one year in Italian pediatric population with type 1 diabetes mellitus: a multicenter retrospective observational study. *Nutr Metab Cardiovasc Dis*, 24(5), 538-546. doi:10.1016/j.numecd.2013.11.004

- Chiavaroli, Derraik, Jalaludin, Albert, Ramkumar, Cutfield, *et al.* (2019). Partial remission in type 1 diabetes and associated factors: Analysis based on the insulin dose-adjusted hemoglobin A1c in children and adolescents from a regional diabetes center, Auckland, New Zealand. *Pediatr Diabetes*, 20(7), 892-900. doi:10.1111/pedi.12881
- Cho, Craig, & Donaghue. (2014). Puberty as an accelerator for diabetes complications. *Pediatr Diabetes*, 15(1), 18-26. doi:10.1111/pedi.12112
- Clarke, Anderson, Pettersson, Cardon, Morris, & Zondervan. (2011). Basic statistical analysis in genetic case-control studies. *Nat. Protocols*, 6(2), 121-133. doi:10.1038/nprot.2010.182
- Clee, Yandell, Schueler, Rabaglia, Richards, Raines, *et al.* (2006). Positional cloning of Sorcs1, a type 2 diabetes quantitative trait locus. *Nat Genet*, 38(6), 688-693. doi:10.1038/ng1796
- Clements, Foster, Maahs, Schatz, Olson, Tsalikian, *et al.* (2016). Hemoglobin A1c (HbA1c) changes over time among adolescent and young adult participants in the T1D exchange clinic registry. *Pediatr Diabetes*, 17(5), 327-336. doi:10.1111/pedi.12295
- Clements, Schwandt, Donaghue, Miller, Lück, Couper, *et al.* (2019). Five heterogeneous HbA1c trajectories from childhood to adulthood in youth with type 1 diabetes from three different continents: A group-based modeling approach. *Pediatr Diabetes*, 20(7), 920-931. doi:10.1111/pedi.12907
- Clerget-Darpoux, Bonaiti-Pellie, & Hochez. (1986). Effects of misspecifying genetic parameters in lod score analysis. *Biometrics*, 42(2), 393-399.
- Cohen, Snieder, Lindsell, Beyan, Hawa, Blinko, *et al.* (2006). Evidence for independent heritability of the glycation gap (glycosylation gap) fraction of HbA1c in nondiabetic twins. *Diabetes Care*, 29(8), 1739-1743. doi:10.2337/dc06-0286
- Comitee. (2007). Consensus statement on the worldwide standardization of the hemoglobin A1C measurement: the American Diabetes Association, European Association for the Study of Diabetes, International Federation of Clinical Chemistry and Laboratory Medicine, and the International Diabetes Federation. *Diabetes Care*, 30(9), 2399-2400. doi:10.2337/dc07-9925
- Comitee. (2009). International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care*, 32(7), 1327-1334. doi:10.2337/dc09-9033
- Cooper, Smyth, Smiles, Plagnol, Walker, Allen, *et al.* (2008). Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat Genet*, 40(12), 1399-1401. doi:10.1038/ng.249
- Cordell. (2009). Detecting gene-gene interactions that underlie human diseases. *Nature reviews. Genetics*, 10(6), 392-404. doi:10.1038/nrg2579
- Coulon, Velasco-Gonzalez, Scribner, Park, Gomez, Vargas, *et al.* (2017). Racial differences in neighborhood disadvantage, inflammation and metabolic control in black and white pediatric type 1 diabetes patients. *Pediatr Diabetes*, 18(2), 120-127. doi:10.1111/pedi.12361
- Dabelea, Rewers, Stafford, Standiford, Lawrence, Saydah, *et al.* (2014). Trends in the prevalence of ketoacidosis at diabetes diagnosis: the SEARCH for diabetes in youth study. *Pediatrics*, 133(4), e938-945. doi:10.1542/peds.2013-2795
- Daly, Vale, Walker, Littlefield, Alberti, & Mathers. (1998). Acute effects on insulin sensitivity and diurnal metabolic profiles of a high-sucrose compared with a high-starch diet. *Am J Clin Nutr*, 67(6), 1186-1196.
- Danne, Philotheou, Goldman, Guo, Ping, Cali, *et al.* (2013). A randomized trial comparing the rate of hypoglycemia--assessed using continuous glucose monitoring--in 125 preschool children with type 1 diabetes treated with insulin glargine or NPH insulin (the PRESCHOOL study). *Pediatr Diabetes*, (8), 593-601. Retrieved from doi:10.1111/pedi.12051

- DCCT. (1994). Effect of intensive diabetes treatment on the development and progression of long-term complications in adolescents with insulin-dependent diabetes mellitus: Diabetes Control and Complications Trial. Diabetes Control and Complications Trial Research Group. *J Pediatr*, *125*(2), 177-188. doi:a56496 [pii]
- de Beaufort, Lange, Swift, Aman, Cameron, Castano, *et al.* (2013). Metabolic outcomes in young children with type 1 diabetes differ between treatment centers: the Hvidoere Study in Young Children 2009. *Pediatr Diabetes*, *14*(6), 422-428. doi:10.1111/j.1399-5448.2012.00922.x
- de Bock, Gunn, Holt, Derraik, Reed, Cutfield, *et al.* (2012). Impact of insulin pumps on glycaemic control in a pump-naive paediatric regional population. *J Paediatr Child Health*, *48*(3), 247-252. doi:10.1111/j.1440-1754.2011.02245.x [doi]
- De, Bush, & Moore. (2014). Bioinformatics challenges in genome-wide association studies (GWAS). *Methods Mol Biol*, *1168*, 63-81. doi:10.1007/978-1-4939-0847-9\_5
- Deeb, Akle, Abdulrahman, Suwaidi, Awad, & Remeithi. (2019). Using insulin pump with a remote-control system in young patients with diabetes improves glycemic control and enhances patient satisfaction. *Clinical diabetes and endocrinology*, *5*, 7-7. doi:10.1186/s40842-019-0081-z
- Deladoey, Henderson, & Geoffroy. (2013). Linear association between household income and metabolic control in children with insulin-dependent diabetes mellitus despite free access to health care. *J Clin Endocrinol Metab*, *98*(5), E882-885. doi:10.1210/jc.2013-1212
- Demirel, Tepe, Esen, Buber, & Boztepe. (2013). Individual and familial factors associated with metabolic control in children with type 1 diabetes. *Pediatrics international : official journal of the Japan Pediatric Society*, *55*(6), 710-713. doi:10.1111/ped.12183
- DeSalvo, Miller, Hermann, Maahs, Hofer, Clements, *et al.* (2018). Continuous glucose monitoring and glycemic control among youth with type 1 diabetes: International comparison from the T1D Exchange and DPV Initiative. *Pediatr Diabetes*, *19*(7), 1271-1275. doi:10.1111/pedi.12711
- DeWan. (2018). Gene-Gene and Gene-Environment Interactions. *Methods Mol Biol*, *1793*, 89-110. doi:10.1007/978-1-4939-7868-7\_7
- Dizier, Demenais, & Mathieu. (2017). Gain of power of the general regression model compared to Cochran-Armitage Trend tests: simulation study and application to bipolar disorder. *BMC Genet*, *18*(1), 24. doi:10.1186/s12863-017-0486-6
- Dłużniak-Gołaska, Panczyk, Szypowska, Sińska, & Szostak-Węgierek. (2019). Interactive Nutrition Education Is More Effective in Terms of Improved Levels of Glycated Hemoglobin in Adolescent Patients with Poorly Controlled Type 1 Diabetes - A Randomized Study. *Diabetes, metabolic syndrome and obesity : targets and therapy*, *12*, 2619-2631. doi:10.2147/DMSO.S229246
- Drew, Berg, King, Verdant, Griffith, Butler, *et al.* (2011). Depleted parental psychological resources as mediators of the association of income with adherence and metabolic control. *Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *25*(5), 751-758. doi:10.1037/a0025259
- Dupuis, Langenberg, Prokopenko, Saxena, Soranzo, Jackson, *et al.* (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*, *42*(2), 105-116. doi:10.1038/ng.520
- Eehalt, Gauger, Blumenstock, Feldhahn, Scheffner, Schweizer, *et al.* (2010). Hemoglobin A1c is a reliable criterion for diagnosing type 1 diabetes in childhood and adolescence. *Pediatr Diabetes*, *11*(7), 446-449. doi:10.1111/j.1399-5448.2009.00633.x
- Eleftheriadou, Didangelos, Pappas, Anastasiou, Vasilopoulos, Zoupas, *et al.* (2019). Improvement of metabolic control after 3-month use of real-time continuous glucose monitoring in patients with

- type 1 diabetes: a multicenter study in Greece. *Hormones (Athens, Greece)*, 10.1007/s42000-42019-00153-42001. doi:10.1007/s42000-019-00153-1
- Eschwege, & Balkau. (2003). [Metabolic syndrome or insulin resistance syndrome. Recent epidemiological data]. *Ann Endocrinol (Paris)*, 64(3 Suppl), S32-36. doi:MDOI-AE-06-2003-64-S3-0003-4266-101019-ART7 [pii]
- Falconer, & Mackay. (1960). Introduction to Quantitative Genetics. *Biometrics*, 38(4).
- Fan, Jung, & Jin. (2006). High-resolution association mapping of quantitative trait loci: a population-based approach. *Genetics*, 172(1), 663-686. doi:10.1534/genetics.105.046417
- Fan, & Xiong. (2002). High resolution mapping of quantitative trait loci by linkage disequilibrium analysis. *Eur J Hum Genet*, 10(10), 607-615. doi:10.1038/sj.ejhg.5200843
- Fendler, Baranowska, Mianowska, Szadkowska, & Mlynarski. (2012). Three-year comparison of subcutaneous insulin pump treatment with multi-daily injections on HbA1c, its variability and hospital burden of children with type 1 diabetes. *Acta diabetologica*, 49(5), 363-370. doi:10.1007/s00592-011-0332-7
- Ferrito, Predieri, Pjetraj, Alessandrelli, Pagnini, Iannilli, *et al.* (2019). Weekend-Based Parent-Group Intervention to Reduce Stress in Parents of Children and Adolescents with Type 1 Diabetes: A Pilot Study. *Journal of diabetes research*, 2019, 7935945-7935945. doi:10.1155/2019/7935945
- FID. (2019). International Diabetes Federation. IDF Diabetes Atlas (Atlas du DIABETE de la FID) - Brussels, Belgium: International Diabetes Federation, 2019. *Neuvième édition*.
- Fisher. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399-433.
- Florez. (2010). A genome-wide association study of treated A1C: a genetic needle in an environmental haystack? *Diabetes*, 59(2), 332-334. doi:10.2337/db09-1636
- Fonolleda, Murillo, Vazquez, Bel, & Vives-Pi. (2017). Remission Phase in Paediatric Type 1 Diabetes: New Understanding and Emerging Biomarkers. *Horm Res Paediatr*, 88(5), 307-315. doi:10.1159/000479030
- Fortenberry, Berg, King, Stump, Butler, Pham, *et al.* (2014). Longitudinal trajectories of illness perceptions among adolescents with type 1 diabetes. *J Pediatr Psychol*, 39(7), 687-696. doi:10.1093/jpepsy/jsu043
- Fortins, Lacerda, Silverio, do Carmo, Ferreira, Felizardo, *et al.* (2019). Predictor factors of glycemic control in children and adolescents with type 1 diabetes mellitus treated at a referral service in Rio de Janeiro, Brazil. *Diabetes Res Clin Pract*, 154, 138-145. doi:10.1016/j.diabres.2019.05.027
- Franklin, Aulchenko, Huffman, Vitart, Hayward, Polasek, *et al.* (2010). The TCF7L2 diabetes risk variant is associated with HbA(1)(C) levels: a genome-wide association meta-analysis. *Ann Hum Genet*, 74(6), 471-478. doi:10.1111/j.1469-1809.2010.00607.x
- Frazer, Ballinger, Cox, Hinds, Stuve, Gibbs, *et al.* (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature*, 449(7164), 851-861. doi:10.1038/nature06258
- Frazer, Murray, Schork, & Topol. (2009). Human genetic variation and its contribution to complex traits. *Nat Rev Genet*, 10(4), 241-251. doi:10.1038/nrg2554
- Fredheim, Johannesen, Johansen, Lyngsoe, Rida, Andersen, *et al.* (2013). Diabetic ketoacidosis at the onset of type 1 diabetes is associated with future HbA1c levels. *Diabetologia*, 56(5), 995-1003. doi:10.1007/s00125-013-2850-z
- Freidlin, Zheng, Li, & Gastwirth. (2002). Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum Hered*, 53(3), 146-152. doi:64976



- Frohlich-Reiterer, Rosenbauer, Bechtold-Dalla Pozza, Hofer, Schober, & Holl. (2014). Predictors of increasing BMI during the course of diabetes in children and adolescents with type 1 diabetes: data from the German/Austrian DPV multicentre survey. *Arch Dis Child*, 99(8), 738-743. doi:10.1136/archdischild-2013-304237
- Gallagher, Le Roith, & Bloomgarden. (2009). Review of hemoglobin A1c in the management of diabetes. *Journal of Diabetes*, 1(1), 9-17. doi:10.1111/j.1753-0407.2009.00009.x
- Galler, Lindau, Ernert, Thalemann, & Raile. (2011). Associations between media consumption habits, physical activity, socioeconomic status, and glycemic control in children, adolescents, and young adults with type 1 diabetes. *Diabetes Care*, 34(11), 2356-2359. doi:10.2337/dc11-0838
- Gao. (2011). Multiple testing corrections for imputed SNPs. *Genetic epidemiology*, 35(3), 154-158. doi:10.1002/gepi.20563
- Gao, Starmer, & Martin. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol*, 32(4), 361-369. doi:10.1002/gepi.20310
- Gao, Wu, Li, Li, Li, & Yang. (2013). Forward LASSO analysis for high-order interactions in genome-wide association study. *Briefings in Bioinformatics*, 15(4), 552-561. doi:10.1093/bib/bbt037
- Garner, Tatu, Best, Creary, & Thein. (2002). Evidence of genetic interaction between the beta-globin complex and chromosome 8q in the expression of fetal hemoglobin. *Am J Hum Genet*, 70(3), 793-799. doi:10.1086/339248
- Gauderman. (2002). Sample size requirements for matched case-control studies of gene-environment interaction. *Stat Med*, 21(1), 35-50.
- Genin, & Clerget-Darpoux. (2015). The missing heritability paradigm: a dramatic resurgence of the GIGO syndrome in genetics. *Hum Hered*, 79(1), 1-4. doi:10.1159/000370327
- George, & Erasmus. (2018). Haemoglobin A1c or Glycated Albumin for Diagnosis and Monitoring Diabetes: An African Perspective. *Indian Journal of Clinical Biochemistry*, 33(3), 255-261. doi:10.1007/s12291-018-0755-9
- Gerber, Locher, Schmid, Spinass, & Lehmann. (2013). Smoking is associated with impaired long-term glucose metabolism in patients with type 1 diabetes mellitus. *Nutr Metab Cardiovasc Dis*, 23(2), 102-108. doi:10.1016/j.numecd.2011.08.007
- Geremia, Fornari, & Tschiedel. (2019). Comparison of the effect of a compact vs a conventional, long-term education program on metabolic control in children and adolescents with type 1 diabetes: A pilot, randomized clinical trial. *Pediatr Diabetes*, 20(6), 778-784. doi:10.1111/pedi.12879
- Gerstl, Rabl, Rosenbauer, Gröbe, Hofer, Krause, et al. (2008). Metabolic control as reflected by HbA1c in children, adolescents and young adults with type-1 diabetes mellitus: combined longitudinal analysis including 27,035 patients from 207 centers in Germany and Austria during the last decade. *European journal of pediatrics*, 167(4), 447-453. doi:10.1007/s00431-007-0586-9
- Gloaguen, Bendelac, Nicolino, Julier, & Mathieu. (2018). A systematic review of non-genetic predictors and genetic factors of glycosylated haemoglobin level in type 1 diabetes patients after the first year of diabetes. *Diabetes Metab Res Rev*, e3051. doi:10.1002/dmrr.3051
- Gloaguen, Dizier, Boissel, Rocheleau, Canouil, Froguel, et al. (2019). General regression model: A "model-free" association test for quantitative traits allowing to test for the underlying genetic model. *Annals of Human Genetics*, n/a(n/a). doi:10.1111/ahg.12372
- Goksen, Atik Altinok, Ozen, Demir, & Darcan. (2014). Effects of carbohydrate counting method on metabolic control in children with type 1 diabetes mellitus. *Journal of clinical research in pediatric endocrinology*, 6(2), 74-78. doi:10.4274/Jcrpe.1191

- Goldenberg, Cheng, Punthakee, & Clement. (2011). Use of Glycated Hemoglobin (A1C) in the Diagnosis of Type 2 Diabetes Mellitus in Adults. *Canadian Journal of Diabetes*, 35(3), 247-249. doi:10.1016/S1499-2671(11)53006-0
- Grabert, Schweiggert, & Holl. (2002). A framework for diabetes documentation and quality management in Germany: 10 years of experience with DPV. *Comput Methods Programs Biomed*, 69(2), 115-121. doi:10.1016/s0169-2607(02)00035-4
- Granhall, Park, Fakhrai-Rad, & Luthman. (2006). High-Resolution Quantitative Trait Locus Analysis Reveals Multiple Diabetes Susceptibility Loci Mapped to Intervals <math>\approx</math>800 kb in the Species-Conserved Niddm1i of the GK Rat. *Genetics*, 174(3), 1565-1572. doi:10.1534/genetics.106.062208
- Grant, Qu, Bradfield, Marchand, Kim, Glessner, *et al.* (2009). Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. *Diabetes*, 58(1), 290-295. doi:10.2337/db08-1022
- Gusareva, & Van Steen. (2014). Practical aspects of genome-wide association interaction analysis. *Human genetics*, 133(11), 1343-1358. doi:10.1007/s00439-014-1480-y
- Haas, Persson, Hagström Toft, Rathsmann, Brorsson, & Lindholm Olander. (2019). Treatment satisfaction correlated with glycaemic control and burden of diabetes in Swedish adolescents with type 1 diabetes. *Acta paediatrica (Oslo, Norway : 1992)*, 10.1111/apa.14991. doi:10.1111/apa.14991
- Hachiya, Komaki, Hasegawa, Ohmomo, Tanno, Hozawa, *et al.* (2017). Genome-wide meta-analysis in Japanese populations identifies novel variants at the TMC6-TMC8 and SIX3-SIX2 loci associated with HbA1c. *Sci Rep*, 7(1), 16147. doi:10.1038/s41598-017-16493-0
- Hakonarson, Grant, Bradfield, Marchand, Kim, Glessner, *et al.* (2007). A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature*, 448(7153), 591-594. doi:10.1038/nature06010
- Hakonarson, Qu, Bradfield, Marchand, Kim, Glessner, *et al.* (2008). A novel susceptibility locus for type 1 diabetes on Chr12q13 identified by a genome-wide association study. *Diabetes*, 57(4), 1143-1146. doi:10.2337/db07-1305
- Hanberger, Akesson, & Samuelsson. (2014). Glycated haemoglobin variations in paediatric type 1 diabetes: the impact of season, gender and age. *Acta Paediatr*, 103(4), 398-403. doi:10.1111/apa.12530
- Hanberger, Samuelsson, Lindblad, & Ludvigsson. (2008). A1C in children and adolescents with diabetes in relation to certain clinical parameters: the Swedish Childhood Diabetes Registry SWEDIABKIDS. *Diabetes Care*, 31(5), 927-929. doi:10.2337/dc07-1863
- Hao, Gitelman, DiMeglio, Boulware, Greenbaum, & Type 1 Diabetes TrialNet Study. (2016). Fall in C-Peptide During First 4 Years From Diagnosis of Type 1 Diabetes: Variable Relation to Age, HbA1c, and Insulin Dose. *Diabetes Care*, 39(10), 1664-1670. doi:10.2337/dc16-0360
- HapMap. (2003). The International HapMap Project. *Nature*, 426(6968), 789-796. doi:10.1038/nature02168
- HapMap. (2005). A haplotype map of the human genome. *Nature*, 437(7063), 1299-1320. doi:10.1038/nature04226
- Harris, Taylor, Thielke, Payne, Gonzalez, & Conde. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377-381. doi:10.1016/j.jbi.2008.08.010

- Harris, Wysocki, Sadler, Wilkinson, Harvey, Buckloh, *et al.* (2000). Validation of a structured interview for the assessment of diabetes self-management. *Diabetes Care*, 23(9), 1301-1304. doi:10.2337/diacare.23.9.1301
- Haseman, & Elston. (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet*, 2(1), 3-19.
- He, Li, Zhang, Su, Hou, Chen, *et al.* (2015). Multi-breed genome-wide association study reveals novel loci associated with the weight of internal organs. *Genetics, selection, evolution : GSE*, 47, 87-87. doi:10.1186/s12711-015-0168-7
- Healey, Gray, Chae, Taylor, Lawrence, Reith, *et al.* (2018). The role of parent and child self-regulation in children's glycemic control. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, 37(4), 326-333. doi:10.1037/hea0000554
- Hebbar, Elkum, Alkayal, John, Thanaraj, & Alsmadi. (2017). Genetic risk variants for metabolic traits in Arab populations. *Sci Rep*, 7, 40988. doi:10.1038/srep40988
- Helgeson, Becker, Escobar, & Siminerio. (2012). Families with children with diabetes: implications of parent stress for parent and child health. *J Pediatr Psychol*, 37(4), 467-478. doi:10.1093/jpepsy/jsr110
- Helgeson, Vaughn, Seltman, Orchard, Libman, & Becker. (2018). Featured Article: Trajectories of Glycemic Control Over Adolescence and Emerging Adulthood: An 11-Year Longitudinal Study of Youth With Type 1 Diabetes. *Journal of pediatric psychology*, 43(1), 8-18. doi:10.1093/jpepsy/jsx083
- Herman. (2016). Are There Clinical Implications of Racial Differences in HbA1c? Yes, to Not Consider Can Do Great Harm! *Diabetes Care*, 39(8), 1458-1461. doi:10.2337/dc15-2686
- Herman, & Fajans. (2010). Hemoglobin A1c for the diagnosis of diabetes: practical considerations. *Pol Arch Med Wewn*, 120(1-2), 37-40.
- Hermann, Miller, Hofer, Clements, Karges, Foster, *et al.* (2019). The Transatlantic HbA(1c) gap: differences in glycaemic control across the lifespan between people included in the US T1D Exchange Registry and those included in the German/Austrian DPV registry. *Diabetic medicine : a journal of the British Diabetic Association*, 10.1111/dme.14148. doi:10.1111/dme.14148
- Hertel, Johansson, Raeder, Platou, Midthjell, Hveem, *et al.* (2011). Evaluation of four novel genetic variants affecting hemoglobin A1c levels in a population-based type 2 diabetes cohort (the HUNT2 study). *BMC medical genetics*, 12, 20. doi:10.1186/1471-2350-12-20
- Hill, & Robertson. (1968). The effects of inbreeding at loci with heterozygote advantage. *Genetics*, 60(3), 615-628.
- Hilliard, Holmes, Chen, Maher, Robinson, & Streisand. (2013). Disentangling the roles of parental monitoring and family conflict in adolescents' management of type 1 diabetes. *Health Psychol*, 32(4), 388-396. doi:10.1037/a0027811
- Hilliard, Rohan, Carle, Pendley, Delamater, & Drotar. (2011). Fathers' involvement in preadolescents' diabetes adherence and glycemic control. *J Pediatr Psychol*, 36(8), 911-922. doi:10.1093/jpepsy/jsr020
- Hilmi, Pasternak, Friger, Loewenthal, Haim, & Hershkovitz. (2013). Ethnic differences in glycemic control and diabetic ketoacidosis rate among children with diabetes mellitus type 1 in the Negev area. *Isr Med Assoc J*, 15(6), 267-270.
- Hirschhorn, & Daly. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, 6(2), 95-108. doi:10.1038/nrg1521

- Hofer, Raile, Frohlich-Reiterer, Kapellen, Dost, Rosenbauer, *et al.* (2014). Tracking of metabolic control from childhood to young adulthood in type 1 diabetes. *J Pediatr*, *165*(5), 956-961 e951-952. doi:10.1016/j.jpeds.2014.07.001
- Holmes, Chen, Streisand, Marschall, Souter, Swift, *et al.* (2006). Predictors of youth diabetes care behaviors and metabolic control: a structural equation modeling approach. *J Pediatr Psychol*, *31*(8), 770-784. doi:10.1093/jpepsy/jsj083
- Hothorn, & Hothorn. (2009). Order-restricted scores test for the evaluation of population-based case-control studies when the genetic model is unknown. *Biom J*, *51*(4), 659-669. doi:10.1002/bimj.200800203
- Huang, Ellinghaus, Franke, Howie, & Li. (2012). 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. *Eur J Hum Genet*, *20*(7), 801-805. doi:10.1038/ejhg.2012.3
- Hughes, Berg, & Wiebe. (2012). Emotional processing and self-control in adolescents with type 1 diabetes. *J Pediatr Psychol*, *37*(8), 925-934. doi:10.1093/jpepsy/jss062 [doi]
- Iannotti, Nansel, Schneider, Haynie, Simons-Morton, Sobel, *et al.* (2006). Assessing regimen adherence of adolescents with type 1 diabetes. *Diabetes Care*, *29*(10), 2263-2267. doi:10.2337/dc06-0685
- Jacobson, Paterson, Ryan, Cleary, Waberski, Weinger, *et al.* (2010). The associations of apolipoprotein E and angiotensin-converting enzyme polymorphisms and cognitive function in Type 1 diabetes based on an 18-year follow-up of the DCCT cohort. *Diabet Med*, *27*(1), 15-22. doi:10.1111/j.1464-5491.2009.02885.x
- Jesic, Sajic, Bogicevic, Buljugic, & Maglajlic. (2013). The effect of metabolic and hormonal parameters on microalbuminuria in adolescents with type 1 diabetes mellitus. *Srp Arh Celok Lek*, *141*(5-6), 315-319.
- Jinno, Urakami, Horikawa, Kawamura, Kikuchi, Kikuchi, *et al.* (2012). Usefulness of insulin detemir in Japanese children with type 1 diabetes. *Pediatr Int*, *54*(6), 773-779. doi:10.1111/j.1442-200X.2012.03687.x
- Johannesen, Svensson, Bergholdt, Eising, Gramstrup, Frandsen, *et al.* (2011). Hypoglycemia, S-ACE and ACE genotypes in a Danish nationwide population of children and adolescents with type 1 diabetes. *Pediatr Diabetes*, *12*(2), 100-106. doi:10.1111/j.1399-5448.2010.00660.x
- Jonckheere. (1954). A Distribution-free k-sample Test against Ordered Alternatives. *Biometrika*, *41*, 133-145.
- Kaiserman, Jung, Benabbad, Karges, Polak, & Rosilio. (2017). 20 Years of insulin lispro in pediatric type 1 diabetes: a review of available evidence. *Pediatr Diabetes*, *18*(2), 81-94. doi:10.1111/pedi.12401
- Kalyva, Abdul-Rasoul, Kehl, Barkai, & Lukacs. (2016). A cross-cultural study on perceived health-related quality of life in children and adolescents with type 1 diabetes mellitus. *J Diabetes Complications*, *30*(3), 482-487. doi:10.1016/j.jdiacomp.2015.12.021
- Kan, & Dozy. (1978). Antenatal diagnosis of sickle-cell anaemia by D.N.A. analysis of amniotic-fluid cells. *Lancet*, *2*(8096), 910-912. doi:10.1016/s0140-6736(78)91629-x
- Kanai, Akiyama, Takahashi, Matoba, Momozawa, Ikeda, *et al.* (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nature Genetics*, *50*(3), 390-400. doi:10.1038/s41588-018-0047-6
- Karczewski, Weisburd, Thomas, Solomonson, Ruderfer, Kavanagh, *et al.* (2017). The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res*, *45*(D1), D840-D845. doi:10.1093/nar/gkw971

- Karges, Schwandt, Heidtmann, Kordonouri, Binder, Schierloh, *et al.* (2017). Association of Insulin Pump Therapy vs Insulin Injection Therapy With Severe Hypoglycemia, Ketoacidosis, and Glycemic Control Among Children, Adolescents, and Young Adults With Type 1 Diabetes. *JAMA*, *318*(14), 1358-1366. doi:10.1001/jama.2017.13994
- Katsarou, Gudbjörnsdottir, Rawshani, Dabelea, Bonifacio, Anderson, *et al.* (2017). Type 1 diabetes mellitus. *Nature Reviews Disease Primers*, *3*, 17016. doi:10.1038/nrdp.2017.16
- Katz, Mehta, Nansel, Quinn, Lipsky, & Laffel. (2014). Associations of nutrient intake with glycemic control in youth with type 1 diabetes: differences by insulin regimen. *Diabetes Technol Ther*, *16*(8), 512-518. doi:10.1089/dia.2013.0389
- Kaur, Mirza, Brorsson, Floyel, Storling, Mortensen, *et al.* (2016). The genetic and regulatory architecture of ERBB3-type 1 diabetes susceptibility locus. *Mol Cell Endocrinol*, *419*, 83-91. doi:10.1016/j.mce.2015.10.002
- Kawasaki. (2014). Type 1 diabetes and autoimmunity. *Clin Pediatr Endocrinol*, *23*(4), 99-105. doi:10.1297/cpe.23.99
- Kebede, Oler, Gregg, Balloon, Johnson, Mitok, *et al.* (2014). SORCS1 is necessary for normal insulin secretory granule biogenesis in metabolically stressed beta cells. *J Clin Invest*, *124*(10), 4240-4256. doi:10.1172/JCI74072
- Kim, Elmi, Henderson, Cogen, & Kaplowitz. (2012). Characteristics of children with type 1 diabetes and persistent suboptimal glycemic control. *J Clin Res Pediatr Endocrinol*, *4*(2), 82-88. doi:10.4274/jcrpe.663 [doi]
- King, Berg, Butner, Drew, Foster, Donaldson, *et al.* (2012). Longitudinal trajectories of metabolic control across adolescence: associations with parental involvement, adolescents' psychosocial maturity, and health care utilization. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, *50*(5), 491-496. doi:10.1016/j.jadohealth.2011.09.007
- Kitsche, Ritz, & Hothorn. (2016). A General Framework for the Evaluation of Genetic Association Studies Using Multiple Marginal Models. *Hum Hered*, *81*(3), 150-172. doi:10.1159/000448477
- Klee, Bellanné-Chantelot, Depret, Llano, Paget, & Nicolino. (2012). A novel ABCC8 mutation illustrates the variability of the diabetes phenotypes associated with a single mutation. *Diabetes Metab*, *38*(2), 179-182. doi:10.1016/j.diabet.2011.12.001
- Kleinberger, & Pollin. (2015). Personalized medicine in diabetes mellitus: current opportunities and future prospects. *Ann N Y Acad Sci*, *1346*(1), 45-56. doi:10.1111/nyas.12757
- Knox, Quirk, Glazebrook, Randell, & Blake. (2019). Impact of technology-based interventions for children and young people with type 1 diabetes on key diabetes self-management behaviours and prerequisites: a systematic review. *BMC Endocrine Disorders*, *19*(1), 7. doi:10.1186/s12902-018-0331-6
- Konietschke, Libiger, & Hothorn. (2012). Nonparametric evaluation of quantitative traits in population-based association studies when the genetic model is unknown. *PLoS One*, *7*(2), e31242. doi:10.1371/journal.pone.0031242
- Konrad, Datz, Engelsberger, Grulich-Henn, Hoertenhuber, Knauth, *et al.* (2015). Current use of metformin in addition to insulin in pediatric patients with type 1 diabetes mellitus: an analysis based on a large diabetes registry in Germany and Austria. *Pediatr Diabetes*, *16*(7), 529-537. doi:10.1111/pedi.12203
- Korolija, Hadzija, Medvidovic, Pavkovic, Kapitanovic, Renar, *et al.* (2010). Genetic evaluation of the TNF-alpha -238G>A and -308G>A promoter polymorphisms in Croatian patients with type I diabetes. *Hum Immunol*, *71*(12), 1228-1232. doi:10.1016/j.humimm.2010.09.003 [doi]

- Krhac, & Lovrencic. (2019). Update on biomarkers of glycemic control. *World J Diabetes*, *10*(1), 1-15. doi:10.4239/wjd.v10.i1.1
- Kwon, & Goate. (2000). The candidate gene approach. *Alcohol Res Health*, *24*(3), 164-168.
- La Greca, Follansbee, & Skyler. (1990). Developmental and Behavioral Aspects of Diabetes Management in Youngsters. *Children's Health Care*, *19*(3), 132-139. doi:10.1207/s15326888chc1903\_1
- Lachin, McGee, & Palmer. (2014). Impact of C-peptide preservation on metabolic and clinical outcomes in the Diabetes Control and Complications Trial. *Diabetes*, *63*(2), 739-748. doi:10.2337/db13-0881
- Lampert, Sygusch, & Schlack. (2007). [Use of electronic media in adolescence. Results of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS)]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, *50*(5-6), 643-652. doi:10.1007/s00103-007-0225-7
- Lander, & Kruglyak. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*, *11*(3), 241-247. doi:10.1038/ng1195-241
- Landrum, Lee, Riley, Jang, Rubinstein, Church, *et al.* (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*, *42*(Database issue), D980-985. doi:10.1093/nar/gkt1113
- Lange. (1986). A test statistic for the affected-sib-set method. *Annals of Human Genetics*, *50*(3), 283-290. doi:10.1111/j.1469-1809.1986.tb01049.x
- Law, Walsh, Queralt, & Nouwen. (2013). Adolescent and parent diabetes distress in type 1 diabetes: the role of self-efficacy, perceived consequences, family responsibility and adolescent-parent discrepancies. *J Psychosom Res*, *74*(4), 334-339. doi:10.1016/j.jpsychores.2012.12.009
- Lee, Abecasis, Boehnke, & Lin. (2014). Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet*, *95*(1), 5-23. doi:10.1016/j.ajhg.2014.06.009
- Lee, Karrison, Cox, & Im. (2013). Quantitative allelic test--a fast test for very large association studies. *Genet Epidemiol*, *37*(8), 831-839. doi:10.1002/gepi.21768
- Lek, Karczewski, Minikel, Samocha, Banks, Fennell, *et al.* (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285-291. doi:10.1038/nature19057
- Lettre, Lange, & Hirschhorn. (2007). Genetic model testing and statistical power in population-based association studies of quantitative traits. *Genet Epidemiol*, *31*(4), 358-362. doi:10.1002/gepi.20217
- Lewontin. (1964). The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. *Genetics*, *49*(1), 49-67.
- Lewontin. (1988). On measures of gametic disequilibrium. *Genetics*, *120*(3), 849-852.
- Li, Li, & Guan. (2008). Evaluation of coverage variation of SNP chips for genome-wide association studies. *Eur J Hum Genet*, *16*(5), 635-643. doi:10.1038/sj.ejhg.5202007
- Li, Li, Zhang, & Snyder. (2019). Gene-Environment Interaction in the Era of Precision Medicine. *Cell*, *177*(1), 38-44. doi:10.1016/j.cell.2019.03.004
- Li, Li, Zheng, Gao, & Yu. (2013). Rank-based robust tests for quantitative-trait genetic association studies. *Genet Epidemiol*, *37*(4), 358-365. doi:10.1002/gepi.21723
- Libman, Miller, DiMeglio, Bethin, Katz, Shah, *et al.* (2015). Effect of Metformin Added to Insulin on Glycemic Control Among Overweight/Obese Adolescents With Type 1 Diabetes: A Randomized Clinical Trial. *JAMA*, *314*(21), 2241-2250. doi:10.1001/jama.2015.16174

- Little, Rohlfing, & Sacks. (2011). Status of hemoglobin A1c measurement and goals for improvement: from chaos to order for improving diabetes care. *Clin Chem*, 57(2), 205-214. doi:10.1373/clinchem.2010.148841
- Loley, Konig, Hothorn, & Ziegler. (2013). A unifying framework for robust association testing, estimation, and genetic model selection using the generalized linear model. *Eur J Hum Genet*, 21(12), 1442-1448. doi:10.1038/ejhg.2013.62
- Loomis, Li, Maruthur, Baldrige, North, Mei, *et al.* (2018). Genome-Wide Association Study of Serum Fructosamine and Glycated Albumin in Adults Without Diagnosed Diabetes: Results From the Atherosclerosis Risk in Communities Study. *Diabetes*, 67(8), 1684-1696. doi:10.2337/db17-1362
- Lukacs, Kiss-Toth, Varga, Soos, Takac, & Barkai. (2013). Benefits of continuous subcutaneous insulin infusion on quality of life. *Int J Technol Assess Health Care*, 29(1), 48-52. doi:10.1017/S0266462312000797
- Lyssenko, Nagorny, Erdos, Wierup, Jonsson, Spegel, *et al.* (2009). Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat Genet*, 41(1), 82-88. doi:10.1038/ng.288
- Maahs, Hermann, DuBose, Miller, Heidtmann, DiMeglio, *et al.* (2014). Contrasting the clinical care and outcomes of 2,622 children with type 1 diabetes less than 6 years of age in the United States T1D Exchange and German/Austrian DPV registries. *Diabetologia*, 57(8), 1578-1585. doi:10.1007/s00125-014-3272-2
- Maahs, Hermann, Holman, Foster, Kapellen, Allgrove, *et al.* (2015). Rates of diabetic ketoacidosis: international comparison with 49,859 pediatric patients with type 1 diabetes from England, Wales, the U.S., Austria, and Germany. *Diabetes Care*, 38(10), 1876-1882. doi:10.2337/dc15-0780
- MacArthur, Bowler, Cerezo, Gil, Hall, Hastings, *et al.* (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*, 45(D1), D896-D901. doi:10.1093/nar/gkw1133
- Machiela, & Chanock. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*, 31(21), 3555-3557. doi:10.1093/bioinformatics/btv402
- Mackey, Struempfl, Powell, Chen, Streisand, & Holmes. (2014). Maternal depressive symptoms and disease care status in youth with type 1 diabetes. *Health Psychol*, 33(8), 783-791. doi:10.1037/hea0000066
- Mahajan, Go, Zhang, Below, Gaulton, Ferreira, *et al.* (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet*, 46(3), 234-244. doi:10.1038/ng.2897
- Mair, Wulaningsih, Jeyam, McGurnaghan, Blackburn, Kennon, *et al.* (2019). Glycaemic control trends in people with type 1 diabetes in Scotland 2004-2016. *Diabetologia*, 62(8), 1375-1384. doi:10.1007/s00125-019-4900-7
- Mameli, Scaramuzza, Ho, Cardona-Hernandez, Suarez-Ortega, & Zuccotti. (2014). A 7-year follow-up retrospective, international, multicenter study of insulin pump therapy in children and adolescents with type 1 diabetes. *Acta Diabetol*, 51(2), 205-210. doi:10.1007/s00592-013-0481-y
- Marchini, Cardon, Phillips, & Donnelly. (2004). The effects of human population structure on large genetic association studies. *Nat Genet*, 36(5), 512-517. doi:10.1038/ng1337
- Marcovecchio, Heywood, Dalton, & Dunger. (2014). The contribution of glycemic control to impaired growth during puberty in young people with type 1 diabetes and microalbuminuria. *Pediatr Diabetes*, 15(4), 303-308. doi:10.1111/peidi.12090 [doi]

- Marees, de Kluiver, Stringer, Vorspan, Curis, Marie-Claire, *et al.* (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int J Methods Psychiatr Res*, 27(2), e1608. doi:10.1002/mpr.1608
- Marissal-Arvy, Helies, Tridon, Moisan, & Mormede. (2014). Quantitative trait Loci influencing abdominal fat deposition and functional variability of the HPA axis in the rat. *Horm Metab Res*, 46(9), 635-643. doi:10.1055/s-0034-1383574
- Marullo, El-Sayed Moustafa, & Prokopenko. (2014). Insights into the genetic susceptibility to type 2 diabetes from genome-wide association studies of glycaemic traits. *Curr Diab Rep*, 14(11), 551. doi:10.1007/s11892-014-0551-8
- Maruthur, Kao, Clark, Brancati, Cheng, Pankow, *et al.* (2011). Does genetic ancestry explain higher values of glycated hemoglobin in African Americans? *Diabetes*, 60(9), 2434-2438. doi:10.2337/db11-0319
- Mathieu, Gillard, & Benhalima. (2017). Insulin analogues in type 1 diabetes mellitus: getting better all the time. *Nat Rev Endocrinol*, 13(7), 385-399. doi:10.1038/nrendo.2017.39
- Mauras, Beck, Xing, Ruedy, Buckingham, Tansey, *et al.* (2012). A randomized clinical trial to assess the efficacy and safety of real-time continuous glucose monitoring in the management of type 1 diabetes in young children aged 4 to <10 years. *Diabetes Care*, 35(2), 204-210. doi:10.2337/dc11-1746
- Mavinkurve, Quinn, & O’Gorman. (2016). Continuous subcutaneous insulin infusion therapy for Type 1 diabetes mellitus in children. *Irish Journal of Medical Science (1971 -)*, 185(2), 335-340. doi:10.1007/s11845-016-1441-7
- Max Andersen, Hougaard, Porksen, Nielsen, Fredheim, Svensson, *et al.* (2014). Partial remission definition: validation based on the insulin dose-adjusted HbA1c (IDAA1C) in 129 Danish children with new-onset type 1 diabetes. *Pediatr Diabetes*, 15(7), 469-476. doi:10.1111/pedi.12208
- Mazarello Paes, Barrett, Dunger, Gevers, Taylor-Robinson, Viner, *et al.* (2020). Factors predicting poor glycemic control in the first two years of childhood onset type 1 diabetes in a cohort from East London, UK: Analyses using mixed effects fractional polynomial models. *Pediatr Diabetes*, 21(2), 288-299. doi:10.1111/pedi.12950
- Mazarello Paes, Charalampopoulos, Edge, Taylor-Robinson, Stephenson, & Amin. (2017). Predictors of glycemic control in the first year of diagnosis of childhood onset type 1 diabetes: A systematic review of quantitative evidence. *Pediatr Diabetes*. doi:10.1111/pedi.12530
- McCarroll, & Altshuler. (2007). Copy-number variation and association studies of human disease. *Nat Genet*.
- McCarthy, Abecasis, Cardon, Goldstein, Little, Ioannidis, *et al.* (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*, 9(5), 356-369. doi:10.1038/nrg2344
- Meigs, Grant, Piccolo, Lopez, Florez, Porneala, *et al.* (2014). Association of African genetic ancestry with fasting glucose and HbA1c levels in non-diabetic individuals: the Boston Area Community Health (BACH) Prediabetes Study. *Diabetologia*, 57(9), 1850-1858. doi:10.1007/s00125-014-3301-1
- Meigs, Manning, Fox, Florez, Liu, Cupples, *et al.* (2007). Genome-wide association with diabetes-related traits in the Framingham Heart Study. *BMC medical genetics*, 8 Suppl 1, S16. Retrieved from doi:10.1186/1471-2350-8-s1-s16



- Meigs, Panhuysen, Myers, Wilson, & Cupples. (2002). A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: The Framingham Offspring Study. *Diabetes*, *51*(3), 833-840.
- Mianowska, Kaminska, Fendler, Szadkowska, & Mlynarski. (2014). Bilirubin is an independent factor inversely associated with glycated hemoglobin level in pediatric patients with type 1 diabetes. *Pediatr Diabetes*, *15*(5), 389-393. doi:10.1111/pedi.12102
- Miller, Beck, Foster, & Maahs. (2020). HbA1c Levels in Type 1 Diabetes from Early Childhood to Older Adults: A Deeper Dive into the influence of technology and socio-economic status on HbA1c in the T1D Exchange Clinic Registry Findings. *Diabetes technology & therapeutics*, 10.1089/dia.2019.0393. doi:10.1089/dia.2019.0393
- Miller, Foster, Beck, Bergenstal, DuBose, DiMeglio, *et al.* (2015). Current state of type 1 diabetes treatment in the U.S.: updated data from the T1D Exchange clinic registry. *Diabetes Care*, *38*(6), 971-978. doi:10.2337/dc15-0078
- Mills, Avery, McCarthy, Hattersley, Levy, Hitman, *et al.* (2004). Heritability estimates for beta cell function and features of the insulin resistance syndrome in UK families with an increased susceptibility to type 2 diabetes. *Diabetologia*, *47*(4), 732-738. doi:10.1007/s00125-004-1338-2
- Mönkemöller, Müller-Godeffroy, Lilienthal, Heidtmann, Becker, Feldhahn, *et al.* (2019). The association between socio-economic status and diabetes care and outcome in children with diabetes type 1 in Germany: The DIAS study (diabetes and social disparities). *Pediatr Diabetes*, *20*(5), 637-644. doi:10.1111/pedi.12847
- Moon, Louie, Jain, Sofer, Schurmann, Below, *et al.* (2019). A Genome-Wide Association Study Identifies Blood Disorder-Related Variants Influencing Hemoglobin A(1c) With Implications for Glycemic Status in U.S. Hispanics/Latinos. *Diabetes Care*, *42*(9), 1784-1791. doi:10.2337/dc19-0168
- Moosavi, Seguin, & Polychronakos. (2017). Effect of autoimmunity risk loci on the honeymoon phase in type 1 diabetes. *Pediatr Diabetes*, *18*(6), 459-462. doi:10.1111/pedi.12421
- Mortensen, Hougaard, Swift, Hansen, Holl, Hoey, *et al.* (2009). New definition for the partial remission period in children and adolescents with type 1 diabetes. *Diabetes Care*, *32*(8), 1384-1390. doi:10.2337/dc08-1987
- Morton. (1955). Sequential tests for the detection of linkage. *Am J Hum Genet*, *7*(3), 277-318.
- Mosaad, Elsharkawy, & El-Deek. (2012). Association of CTLA-4 (+49A/G) gene polymorphism with type 1 diabetes mellitus in Egyptian children. *Immunol Invest*, *41*(1), 28-37. doi:10.3109/08820139.2011.579215
- Musha, Mochizuki, Kikuchi, Akatsuka, Ohtake, Kobayashi, *et al.* (2018). Estimation of glycaemic control in the past month using ratio of glycated albumin to HbA(1c). *Diabetic medicine : a journal of the British Diabetic Association*, *35*(7), 855-861. doi:10.1111/dme.13640
- Must, & Anderson. (2006). Body mass index in children and adolescents: considerations for population-based applications. *Int J Obes (Lond)*, *30*(4), 590-594. doi:10.1038/sj.ijo.0803300
- Nadeau, Chow, Alam, Lindquist, Campbell, McFann, *et al.* (2015). Effects of low dose metformin in adolescents with type I diabetes mellitus: a randomized, double-blinded placebo-controlled study. *Pediatr Diabetes*, *16*(3), 196-203. doi:10.1111/pedi.12140
- Nagl, Hermann, Plamper, Schröder, Dost, Kordonouri, *et al.* (2017). Factors contributing to partial remission in type 1 diabetes: analysis based on the insulin dose-adjusted HbA1c in 3657 children and adolescents from Germany and Austria. *Pediatr Diabetes*, *18*(6), 428-434. doi:10.1111/pedi.12413

- Nansel, Lipsky, & Iannotti. (2013). Cross-sectional and longitudinal relationships of body mass index with glycemic control in children and adolescents with type 1 diabetes mellitus. *Diabetes Res Clin Pract*, *100*(1), 126-132. doi:10.1016/j.diabres.2012.12.025
- Neylon, O'Connell, Skinner, & Cameron. (2013). Demographic and personal factors associated with metabolic control and self-care in youth with type 1 diabetes: a systematic review. *Diabetes Metab Res Rev*, *29*(4), 257-272. doi:10.1002/dmrr.2392
- Nicolae. (2016). Association Tests for Rare Variants. *Annu Rev Genomics Hum Genet*, *17*, 117-130. doi:10.1146/annurev-genom-083115-022609
- Nielsen, Gaulke, Eriksen, Svensson, & Skipper. (2019). Socioeconomic Inequality in Metabolic Control Among Children With Type 1 Diabetes: A Nationwide Longitudinal Study of 4,079 Danish Children. *Diabetes Care*, *42*(8), 1398-1405. doi:10.2337/dc19-0184
- Nielsen, Porksen, Andersen, Fredheim, Svensson, Hougaard, *et al.* (2011). The PTPN22 C1858T gene variant is associated with proinsulin in new-onset type 1 diabetes. *BMC medical genetics*, *12*, 41. doi:10.1186/1471-2350-12-41
- Nielsen, & Weir. (1999). A classical setting for associations between markers and loci affecting quantitative traits. *Genet Res*, *74*(3), 271-277. doi:10.1017/s0016672399004231
- Nielsen, & Weir. (2001). Association studies under general disease models. *Theor Popul Biol*, *60*(3), 253-263. doi:10.1006/tpbi.2001.1539
- Nilsson, Åkesson, Hanberger, & Samuelsson. (2017). High HbA1c at onset cannot be used as a predictor for future metabolic control for the individual child with type 1 diabetes mellitus. *Pediatr Diabetes*, *18*(8), 848-852. doi:10.1111/pedi.12498
- Nordwall, Abrahamsson, Dhir, Fredrikson, Ludvigsson, & Arnqvist. (2015). Impact of HbA1c, followed from onset of type 1 diabetes, on the development of severe retinopathy and nephropathy: the VISS Study (Vascular Diabetic Complications in Southeast Sweden). *Diabetes Care*, *38*(2), 308-315. doi:10.2337/dc14-1203
- Nordwall, Arnqvist, Bojestig, & Ludvigsson. (2009). Good glycemic control remains crucial in prevention of late diabetic complications--the Linköping Diabetes Complications Study. *Pediatr Diabetes*, *10*(3), 168-176. doi:10.1111/j.1399-5448.2008.00472.x
- Nørgaard, Sukumar, Rafnsson, & Saravanan. (2018). Efficacy and Safety of Rapid-Acting Insulin Analogs in Special Populations with Type 1 Diabetes or Gestational Diabetes: Systematic Review and Meta-Analysis. *Diabetes therapy : research, treatment and education of diabetes and related disorders*, *9*(3), 891-917. doi:10.1007/s13300-018-0411-7
- Nylander, Toivonen, Nasic, Soderstrom, Tindberg, & Fernell. (2013). Children and adolescents with type 1 diabetes and high HbA1c -- a neurodevelopmental perspective. *Acta paediatrica (Oslo, Norway : 1992)*, *102*(4), 410-415. doi:10.1111/apa.12128
- Ogilvie, Fayter, Petticrew, Sowden, Thomas, Whitehead, *et al.* (2008). The harvest plot: a method for synthesising evidence about the differential effects of interventions. *BMC Med Res Methodol*, *8*, 8. doi:10.1186/1471-2288-8-8
- Olsen, Johannesen, Fredheim, & Svensson. (2015). Insulin pump treatment; increasing prevalence, and predictors for better metabolic outcome in Danish children and adolescents with type 1 diabetes. *Pediatr Diabetes*, *16*(4), 256-262. doi:10.1111/pedi.12164
- OMS. (2016). Résumé d'orientation : Rapport mondial sur le diabète.
- Onengut-Gumuscu, Chen, Burren, Cooper, Quinlan, Mychaleckyj, *et al.* (2015). Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet*, *47*(4), 381-386. doi:10.1038/ng.3245

- Ozaki, Ohnishi, Iida, Sekine, Yamada, Tsunoda, *et al.* (2002). Functional SNPs in the lymphotoxin- $\alpha$  gene that are associated with susceptibility to myocardial infarction. *Nature Genetics*, 32(4), 650-654. doi:10.1038/ng1047
- Pare, Chasman, Parker, Nathan, Miletich, Zee, *et al.* (2008). Novel association of HK1 with glycated hemoglobin in a non-diabetic population: a genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. *PLoS Genet*, 4(12), e1000312. doi:10.1371/journal.pgen.1000312
- Paterson, Waggott, Boright, Hosseini, Shen, Sylvestre, *et al.* (2010). A Genome-Wide Association Study Identifies a Novel Major Locus for Glycemic Control in Type 1 Diabetes, as Measured by Both A1C and Glucose. *Diabetes*, 59(2), 539-549.
- Patterson, Gyurus, Rosenbauer, Cinek, Neu, Schober, *et al.* (2012). Trends in childhood type 1 diabetes incidence in Europe during 1989-2008: evidence of non-uniformity over time in rates of increase. *Diabetologia*, 55(8), 2142-2147. doi:10.1007/s00125-012-2571-8
- Penrose. (1953). The general purpose sibpair linkage test. *Ann Eugen*, 18(2), 120-124.
- Peterson, Hanberger, Akesson, Bojestig, Andersson Gare, & Samuelsson. (2014). Improved results in paediatric diabetes care using a quality registry in an improvement collaborative: a case study in Sweden. *PLoS One*, 9(5), e97875. doi:10.1371/journal.pone.0097875
- Petit-Bibal, Rothenbuhler, Lucchini, Aboumrad, Castell, Fur, *et al.* (2015). Decrease in clinical hypoglycemia in young children with type 1 diabetes treated with free-mixed aspart and detemir insulin: An open labeled randomized trial. *Pediatr Diabetes*, (5), 345-353. Retrieved from doi:10.1111/pedi.12156
- Petrone, Spoletini, Zampetti, Capizzi, Zavarella, Osborn, *et al.* (2008). The PTPN22 1858T gene variant in type 1 diabetes is associated with reduced residual beta-cell function and worse metabolic control. *Diabetes Care*, 31(6), 1214-1218. doi:10.2337/dc07-1158
- Piccini, Schwandt, Jefferies, Kordonouri, Limbert, Arslanoglu, *et al.* (2019). Association of diabetic ketoacidosis and HbA1c at onset with year-three HbA1c in children and adolescents with type 1 diabetes: Data from the International SWEET Registry. *Pediatr Diabetes*, 10.1111/pedi.12946. doi:10.1111/pedi.12946
- Piffaretti, Mandereau-Bruno, Guilmin-Crepon, Choleau, Coutant, & Fosse-Edorh. (2019). Trends in childhood type 1 diabetes incidence in France, 2010-2015. *Diabetes Res Clin Pract*, 149, 200-207. doi:10.1016/j.diabres.2018.11.005
- Pilia, Chen, Scuteri, Orru, Albai, Dei, *et al.* (2006). Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet*, 2(8), e132. doi:10.1371/journal.pgen.0020132
- Pociot, & Lernmark. (2016). Genetic risk factors for type 1 diabetes. *Lancet*, 387(10035), 2331-2339. doi:10.1016/S0140-6736(16)30582-7
- Prahalad, Tanenbaum, Hood, & Maahs. (2018). Diabetes technology: improving care, improving patient-reported outcomes and preventing complications in young people with Type 1 diabetes. *Diabet Med*, 35(4), 419-429. doi:10.1111/dme.13588
- Price, Knowles, Fox, Wales, Heller, Eiser, *et al.* (2016). Effectiveness of the Kids in Control of Food (KICK-OFF) structured education course for 11-16 year olds with Type 1 diabetes. *Diabet Med*, 33(2), 192-203. doi:10.1111/dme.12881
- Price, Patterson, Plenge, Weinblatt, Shadick, & Reich. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*, 38(8), 904-909. doi:10.1038/ng1847
- Prins, Kuchenbaecker, Bao, Smart, Zabaneh, Fatemifar, *et al.* (2017). Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. *Sci Rep*, 7(1), 11008. doi:10.1038/s41598-017-10812-1

- Prokopenko, Langenberg, Florez, Saxena, Soranzo, Thorleifsson, *et al.* (2009). Variants in MTNR1B influence fasting glucose levels. *Nat Genet*, *41*(1), 77-81. doi:10.1038/ng.290
- Pulgaron, Sanders, Patino-Fernandez, Wile, Sanchez, Rothman, *et al.* (2014). Glycemic control in young children with diabetes: the role of parental health literacy. *Patient Educ Couns*, *94*(1), 67-70. doi:10.1016/j.pec.2013.09.002
- Purcell, Neale, Todd-Brown, Thomas, Ferreira, Bender, *et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, *81*(3), 559-575. doi:10.1086/519795
- Rabbone, Scaramuzza, Ignaccolo, Tinti, Sicignano, Redaelli, *et al.* (2014). Carbohydrate counting with an automated bolus calculator helps to improve glycaemic control in children with type 1 diabetes using multiple daily injection therapy: an 18-month observational study. *Diabetes Res Clin Pract*, *103*(3), 388-394. doi:10.1016/j.diabres.2013.12.026
- Rahbar, Blumenfeld, & Ranney. (1969). Studies of an unusual hemoglobin in patients with diabetes mellitus. *Biochem Biophys Res Commun*, *36*(5), 838-843. doi:0006-291X(69)90685-8 [pii]
- Rahman, Bennet, Pedersen, de Faire, Svensson, & Magnusson. (2009). Genetic dominance influences blood biomarker levels in a sample of 12,000 Swedish elderly twins. *Twin Res Hum Genet*, *12*(3), 286-294. doi:10.1375/twin.12.3.286
- Rechenberg, Whittemore, & Grey. (2017). Anxiety in Youth With Type 1 Diabetes. *J Pediatr Nurs*, *32*, 64-71. doi:10.1016/j.pedn.2016.08.007
- Redon, Beltrand, Martin, Taupin, Choleau, Morandini, *et al.* (2014). Changes in insulin therapy regimens over 10 yr in children and adolescents with type 1 diabetes attending diabetes camps. *Pediatr Diabetes*, *15*(5), 329-335. doi:10.1111/pedi.12062
- Redondo, Connor, Ruedy, Beck, Kollman, Wood, *et al.* (2014). Pediatric Diabetes Consortium Type 1 Diabetes New Onset (NeOn) Study: factors associated with HbA1c levels one year after diagnosis. *Pediatr Diabetes*, *15*(4), 294-302. doi:10.1111/pedi.12061 [doi]
- Redondo, Libman, Cheng, Kollman, Tosur, Gal, *et al.* (2018). Racial/Ethnic Minority Youth With Recent-Onset Type 1 Diabetes Have Poor Prognostic Factors. *Diabetes Care*, *41*(5), 1017-1024. doi:10.2337/dc17-2335
- Reed, Nunez, Kulp, Qian, Reilly, & Foulkes. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Stat Med*, *34*(28), 3769-3792. doi:10.1002/sim.6605
- Rewers, & Ludvigsson. (2016). Environmental risk factors for type 1 diabetes. *Lancet*, *387*(10035), 2340-2348. doi:10.1016/S0140-6736(16)30507-4
- Rewers, Pillay, de Beaufort, Craig, Hanas, Acerini, *et al.* (2014). ISPAD Clinical Practice Consensus Guidelines 2014. Assessment and monitoring of glycemic control in children and adolescents with diabetes. *Pediatr Diabetes*, *15 Suppl 20*, 102-114. doi:10.1111/pedi.12190
- Risch. (1990). Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet*, *46*(2), 229-241.
- Risch, & Merikangas. (1996). The future of genetic studies of complex human diseases. *Science*, *273*(5281), 1516-1517.
- Ritz, Chatterjee, Garcia-Closas, Gauderman, Pierce, Kraft, *et al.* (2017). Lessons Learned From Past Gene-Environment Interaction Successes. *Am J Epidemiol*, *186*(7), 778-786. doi:10.1093/aje/kwx230
- Roelants, Hauspie, & Hoppenbrouwers. (2009). References for growth and pubertal development from birth to 21 years in Flanders, Belgium. *Ann Hum Biol*, *36*(6), 680-694. doi:10.3109/03014460903049074

- Rogers, Kim, Banerjee, & Lee. (2017). Fluctuations in the incidence of type 1 diabetes in the United States from 2001 to 2015: a longitudinal study. *BMC Med*, *15*(1), 199. doi:10.1186/s12916-017-0958-6
- Rohan, Rausch, Pendley, Delamater, Dolan, Reeves, *et al.* (2014). Identification and prediction of group-based glycemic control trajectories during the transition to adolescence. *Health Psychol*, *33*(10), 1143-1152. doi:10.1037/hea0000025
- Rokamp, Olsen, Faerch, Kristensen, Thorsteinsson, & Pedersen-Bjergaard. (2018). Impact of Genetic Polymorphism in the beta2-Receptor Gene on Risk of Severe Hypoglycemia in Patients With Type 1 Diabetes. *J Clin Endocrinol Metab*, *103*(8), 2901-2908. doi:10.1210/jc.2017-02739
- Rosenbauer, Dost, Karges, Hungele, Stahl, Bachle, *et al.* (2012). Improved metabolic control in children and adolescents with type 1 diabetes: a trend analysis using prospective multicenter data from Germany and Austria. *Diabetes Care*, *35*(1), 80-86. doi:10.2337/dc11-0993 [doi]
- Ryu, & Lee. (2012). Association of glycosylated hemoglobin with the gene encoding CDKAL1 in the Korean Association Resource (KARE) study. *Hum Mutat*, *33*(4), 655-659. doi:10.1002/humu.22040
- Saaddine, Fagot-Campagna, Rolka, Narayan, Geiss, Eberhardt, *et al.* (2002). Distribution of HbA(1c) levels for children and young adults in the U.S.: Third National Health and Nutrition Examination Survey. *Diabetes Care*, *25*(8), 1326-1330.
- Sands, Higgins, Mehta, Nansel, Lipsky, & Laffel. (2013). Associations of youth and parent weight status with reported versus predicted daily energy intake and hemoglobin A1c in youth with type 1 diabetes mellitus. *J Diabetes Sci Technol*, *7*(1), 263-270.
- Sanyoura, Philipson, & Naylor. (2018). Monogenic Diabetes in Children and Adolescents: Recognition and Treatment Options. *Curr Diab Rep*, *18*(8), 58. doi:10.1007/s11892-018-1024-2
- Sazonovs, & Barrett. (2018). Rare-Variant Studies to Complement Genome-Wide Association Studies. *Annu Rev Genomics Hum Genet*. doi:10.1146/annurev-genom-083117-021641
- Schaid, McDonnell, Hebring, Cunningham, & Thibodeau. (2005). Nonparametric tests of association of multiple genes with human disease. *Am J Hum Genet*, *76*(5), 780-793. doi:10.1086/429838
- Scheuing, Wiegand, Bachle, Frohlich-Reiterer, Hahn, Icks, *et al.* (2015). Impact of Maternal Country of Birth on Type-1-Diabetes Therapy and Outcome in 27,643 Children and Adolescents from the DPV Registry. *PLoS One*, *10*(8), e0135178. doi:10.1371/journal.pone.0135178
- Schwandt, Hermann, Rosenbauer, Boettcher, Dunstheimer, Grulich-Henn, *et al.* (2017). Longitudinal Trajectories of Metabolic Control From Childhood to Young Adulthood in Type 1 Diabetes From a Large German/Austrian Registry: A Group-Based Modeling Approach. *Diabetes Care*, *40*(3), 309-316. doi:10.2337/dc16-1625
- Schweyer, & Jellimann. (2015). [Adolescence and diabetes, psychological and physiological aspects]. *Rev Infirm*(216), 37-38. doi:10.1016/j.revinf.2015.10.018
- Segal. (2000). *Entwined Lives*: Penguin Group US.
- Selvin. (2016). Are There Clinical Implications of Racial Differences in HbA1c? A Difference, to Be a Difference, Must Make a Difference. *Diabetes Care*, *39*(8), 1462-1467. doi:10.2337/dc16-0042
- Shah, & Green. (1994). Disease susceptibility genes and the sib-pair method: a review of recent methodology. *Ann Hum Genet*, *58*(Pt 4), 381-395.
- Shalitin, & Phillip. (2012). Which factors predict glycemic control in children diagnosed with type 1 diabetes before 6.5 years of age? *Acta diabetologica*, *49*(5), 355-362. doi:10.1007/s00592-011-0321-x

- Sham, & Purcell. (2014). Statistical power and significance testing in large-scale genetic studies. *Nat Rev Genet*, 15(5), 335-346. doi:10.1038/nrg3706
- Sherry, Ward, Kholodov, Baker, Phan, Smigielski, *et al.* (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*, 29(1), 308-311.
- Sherwani, Khan, Ekhzaimy, Masood, & Sakharkar. (2016). Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. *Biomark Insights*, 11, 95-104. doi:10.4137/BMI.S38440
- Shorer, David, Schoenberg-Taz, Levavi-Lavi, Phillip, & Meyerovitch. (2011). Role of parenting style in achieving metabolic control in adolescents with type 1 diabetes. *Diabetes Care*, 34(8), 1735-1737. doi:10.2337/dc10-1602
- Simon, Sylvestre, Tremblay, & Hamet. (2016). Key Considerations and Methods in the Study of Gene-Environment Interactions. *Am J Hypertens*, 29(8), 891-899. doi:10.1093/ajh/hpw021
- Simonis-Bik, Eekhoff, Diamant, Boomsma, Heine, Dekker, *et al.* (2008). The heritability of HbA1c and fasting blood glucose in different measurement settings. *Twin Res Hum Genet*, 11(6), 597-602. doi:10.1375/twin.11.6.597
- Skol, Scott, Abecasis, & Boehnke. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet*, 38(2), 209-213. doi:10.1038/ng1706
- Slager, & Schaid. (2001). Case-control studies of genetic markers: power and sample size approximations for Armitage's test for trend. *Hum Hered*, 52(3), 149-153. doi:10.1159/000053370
- Slingerland. (2006). Monogenic diabetes in children and young adults: Challenges for researcher, clinician and patient. *Reviews in endocrine & metabolic disorders*, 7(3), 171-185. doi:10.1007/s11154-006-9014-0
- Slominski, Lawrynowicz, Ryba-Stanislawowska, Skrzypkowska, Mysliwska, & Mysliwiec. (2019). CCR5-Delta32 polymorphism is a genetic risk factor associated with dyslipidemia in patients with type 1 diabetes. *Cytokine*, 114, 81-85. doi:10.1016/j.cyto.2018.11.005
- So, & Sham. (2011). Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. *Behav Genet*, 41(5), 768-775. doi:10.1007/s10519-011-9450-9
- Sood, Pendley, Delamater, Rohan, Pulgaron, & Drotar. (2012). Mother-father informant discrepancies regarding diabetes management: associations with diabetes-specific family conflict and glycemic control. *Health Psychol*, 31(5), 571-579. doi:10.1037/a0029006
- Soranzo. (2011). Genetic determinants of variability in glycosylated hemoglobin (HbA(1c)) in humans: review of recent progress and prospects for use in diabetes care. *Curr Diab Rep*, 11(6), 562-569. doi:10.1007/s11892-011-0232-9
- Soranzo, Sanna, Wheeler, Gieger, Radke, Dupuis, *et al.* (2010). Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes*, 59(12), 3229-3239. doi:10.2337/db10-0502 [doi]
- Soulimane, Simon, Herman, Lange, Lee, Colagiuri, *et al.* (2014). HbA1c, fasting and 2 h plasma glucose in current, ex- and never-smokers: a meta-analysis. *Diabetologia*, 57(1), 30-39. doi:10.1007/s00125-013-3058-y
- Souza, Ururahy, Oliveira, Loureiro, Silva, Bortolin, *et al.* (2018). The low-density lipoprotein receptor-related protein 5 (LRP5) 4037C>T polymorphism: candidate for susceptibility to type 1 diabetes mellitus. *Arch Endocrinol Metab*, 62(4), 480-484. doi:10.20945/2359-3997000000057
- Spaans, Kleefstra, Groenier, Bilo, & Brand. (2020). Adherence to insulin pump treatment declines with increasing age in adolescents with type 1 diabetes mellitus. *Acta Paediatr*, 109(1), 134-139. doi:10.1111/apa.14931

- Spoletini, Petrone, Zampetti, Capizzi, Zavarella, Osborn, *et al.* (2007). Low-risk HLA genotype in Type 1 diabetes is associated with less destruction of pancreatic B-cells 12 months after diagnosis. *Diabet Med*, *24*(12), 1487-1490. doi:10.1111/j.1464-5491.2007.02292.x
- Spracklen, Shi, Vadlamudi, Wu, Zou, Raulerson, *et al.* (2018). Identification and functional analysis of glycemic trait loci in the China Health and Nutrition Survey. *PLoS Genet*, *14*(4), e1007275. doi:10.1371/journal.pgen.1007275
- Stahl-Pehe, Strassburger, Castillo, Bachle, Holl, Lange, *et al.* (2014). Quality of life in intensively treated youths with early-onset type 1 diabetes: a population-based survey. *Pediatr Diabetes*, *15*(6), 436-443.
- Sullivan. (2007). Spurious genetic associations. *Biol Psychiatry*, *61*(10), 1121-1126. doi:10.1016/j.biopsych.2006.11.010
- Sun, Malcolm, Wong, Shorr, & Doyle. (2019). Improving Glycemic Control in Adults and Children With Type 1 Diabetes With the Use of Smartphone-Based Mobile Applications: A Systematic Review. *Canadian Journal of Diabetes*, *43*(1), 51-58.e53. doi:10.1016/j.cjcd.2018.03.010
- Sung, Lee, Song, Lee, & Kim. (2015). Genetic and baseline metabolic factors for incident diabetes and HbA(1c) at follow-up: the healthy twin study. *Diabetes Metab Res Rev*, *31*(4), 376-384. doi:10.1002/dmrr.2619
- Svensson, Cerqueira, Kjaersgaard, Lyngsoe, Hertel, Madsen, *et al.* (2016). Danish Registry of Childhood and Adolescent Diabetes. *Clin Epidemiol*, *8*, 679-683. doi:10.2147/clep.S99469
- Syreeni, Sandholm, Cao, Toppila, Maahs, Rewers, *et al.* (2019). Genetic Determinants of Glycated Hemoglobin in Type 1 Diabetes. *Diabetes*, *68*(4), 858-867. doi:10.2337/db18-0573
- Szablewski. (2014). Role of immune system in type 1 diabetes mellitus pathogenesis. *International Immunopharmacology*, *22*(1), 182-191. doi:10.1016/j.intimp.2014.06.033
- Szklarczyk, Franceschini, Wyder, Forslund, Heller, Huerta-Cepas, *et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*, *43*(Database issue), D447-452. doi:10.1093/nar/gku1003
- Szumilas. (2010). Explaining Odds Ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *19*(3), 227-229.
- Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org>
- Tejero, Voruganti, Cai, Cole, Laston, Wenger, *et al.* (2010). Pleiotropic effects on subclasses of HDL, adiposity, and glucose metabolism in adult Alaskan Eskimos. *Am J Hum Biol*, *22*(4), 444-448. doi:10.1002/ajhb.21015
- Teló, Dougher, Volkening, Katz, & Laffel. (2018). Predictors of changing insulin dose requirements and glycaemic control in children, adolescents and young adults with Type 1 diabetes. *Diabetic medicine : a journal of the British Diabetic Association*, *35*(10), 1355-1363. doi:10.1111/dme.13699
- Telo, Volkening, Butler, & Laffel. (2015). Salient characteristics of youth with type 1 diabetes initiating continuous glucose monitoring. *Diabetes Technol Ther*, *17*(6), 373-378. doi:10.1089/dia.2014.0290
- Terpstra. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. . *Indagationes Mathematicae*, *14*, 327-333.
- Thomas, Lipsky, Liu, & Nansel. (2018). Income Relates to Adherence in Youth with Type 1 Diabetes Through Parenting Constructs. *Journal of developmental and behavioral pediatrics : JDBP*, *39*(6), 508-515. doi:10.1097/DBP.0000000000000579

- Todd, Walker, Cooper, Smyth, Downes, Plagnol, *et al.* (2007). Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet*, 39(7), 857-864. doi:10.1038/ng2068
- Tsepilov, Shin, Soranzo, Spector, Prehn, Adamski, *et al.* (2015). Nonadditive Effects of Genes in Human Metabolomics. *Genetics*, 200(3), 707-718. doi:10.1534/genetics.115.175760
- Tsugawa, Mukamal, Davis, Taylor, & Wee. (2012). Should the hemoglobin A1c diagnostic cutoff differ between blacks and whites? A cross-sectional study. *Ann Intern Med*, 157(3), 153-159. doi:10.7326/0003-4819-157-3-201208070-00004
- UKPDS. (1998). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet*, 352(9131), 837-853. doi:S0140673698070196 [pii]
- Ururahy, de Souza, Oliveira, Loureiro, da Silva, Freire-Neto, *et al.* (2015). Association of polymorphisms in IL6 gene promoter region with type 1 diabetes and increased albumin-to-creatinine ratio. *Diabetes Metab Res Rev*, 31(5), 500-506. doi:10.1002/dmrr.2621
- Valenzuela, Smith, Stafford, D'Agostino, Lawrence, Yi-Frazier, *et al.* (2014). Shared decision-making among caregivers and health care providers of youth with type 1 diabetes. *J Clin Psychol Med Settings*, 21(3), 234-243. doi:10.1007/s10880-014-9400-9
- van Esdonk, Tai, Cotterill, Charles, & Hennig. (2017). Prediction of glycaemic control in young children and adolescents with type 1 diabetes mellitus using mixed-effects logistic regression modelling. *PLoS One*, 12(8), e0182181-e0182181. doi:10.1371/journal.pone.0182181
- VanBuecken, & Greenbaum. (2014). Residual C-peptide in type 1 diabetes: what do we really know? *Pediatr Diabetes*, 15(2), 84-90. doi:10.1111/pedi.12135
- Venkataraman, Kao, Thai, Salim, Lee, Heng, *et al.* (2012). Ethnicity modifies the relation between fasting plasma glucose and HbA1c in Indians, Malays and Chinese. *Diabet Med*, 29(7), 911-917. doi:10.1111/j.1464-5491.2012.03599.x
- Vesco, Jedraszko, Garza, & Weissberg-Benchell. (2018). Continuous Glucose Monitoring Associated With Less Diabetes-Specific Emotional Distress and Lower A1c Among Adolescents With Type 1 Diabetes. *Journal of diabetes science and technology*, 12(4), 792-799. doi:10.1177/1932296818766381
- Visscher, Brown, McCarthy, & Yang. (2012). Five years of GWAS discovery. *Am J Hum Genet*, 90(1), 7-24. doi:10.1016/j.ajhg.2011.11.029
- Vogel. (2019). childsd: Data and Methods Around Reference Values in Pediatrics. (Version version 0.7.4). Retrieved from <https://CRAN.R-project.org/package=childsd>
- Wallace, Smyth, Maisuria-Armer, Walker, Todd, & Clayton. (2010). The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. *Nat Genet*, 42(1), 68-71. doi:10.1038/ng.493
- Waller, Johnston, Molyneaux, Brown-Singh, Hatherly, Smith, *et al.* (2013). Glycemic control and blood glucose monitoring over time in a sample of young Australians with type 1 diabetes: the role of personality. *Diabetes Care*, 36(10), 2968-2973. doi:10.2337/dc12-1743
- Wang, & Lu. (2017). Epigenetic Alterations in Cellular Immunity: New Insights into Autoimmune Diseases. *Cell Physiol Biochem*, 41(2), 645-660. doi:10.1159/000457944
- Wang, Lu, & Zhao. (2015). A review of study designs and statistical methods for genomic epidemiology studies using next generation sequencing. *Front Genet*, 6, 149. doi:10.3389/fgene.2015.00149



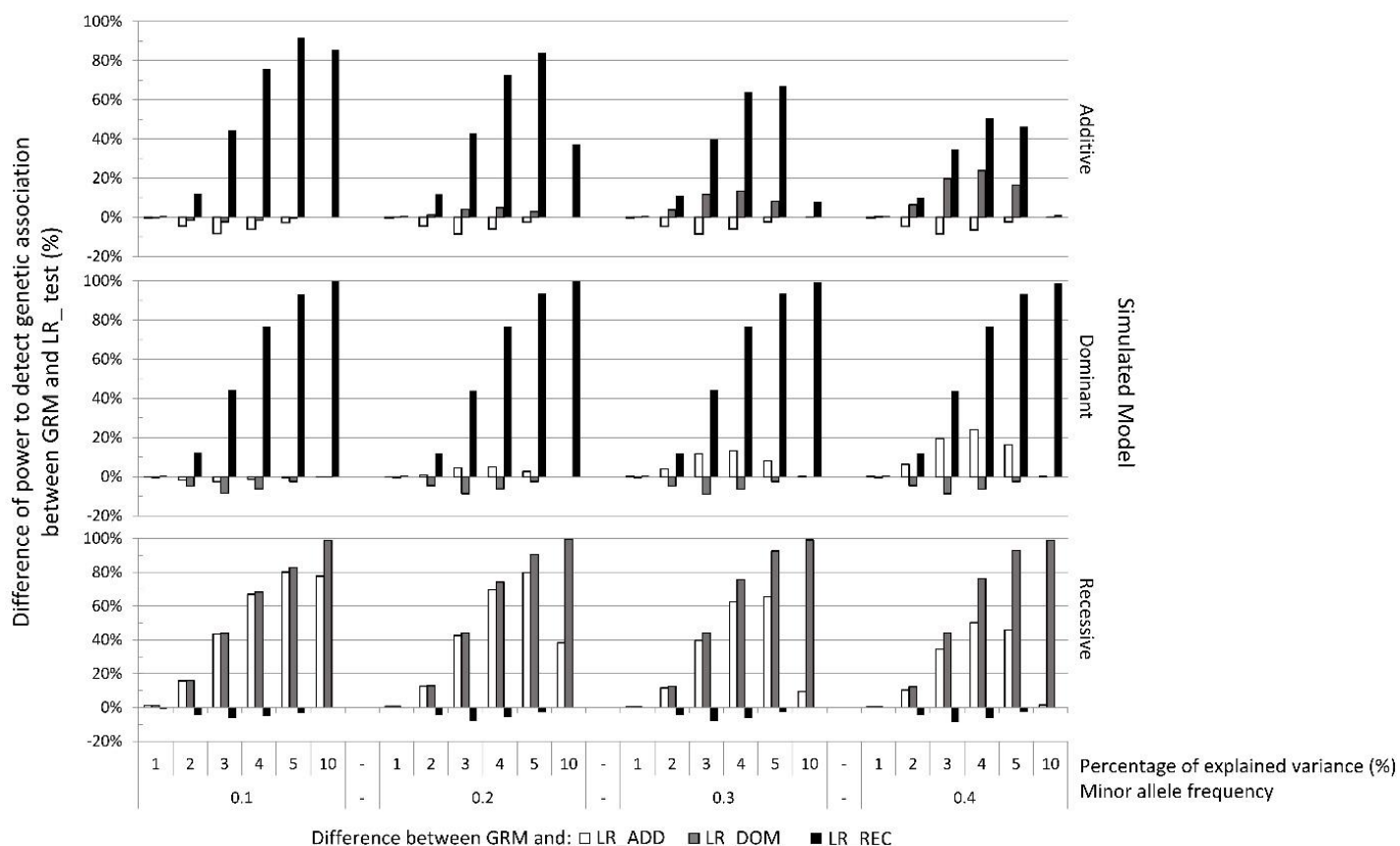
- Wang, Wiebe, & White. (2011). Developmental trajectories of metabolic control among White, Black, and Hispanic youth with type 1 diabetes. *The Journal of pediatrics*, 159(4), 571-576. doi:10.1016/j.jpeds.2011.03.053
- Welter, MacArthur, Morales, Burdett, Hall, Junkins, *et al.* (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*, 42(Database issue), D1001-1006. doi:10.1093/nar/gkt1229
- Wheeler, Leong, Liu, Hivert, Strawbridge, Podmore, *et al.* (2017). Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med*, 14(9), e1002383. doi:10.1371/journal.pmed.1002383
- Whirl-Carrillo, McDonagh, Hebert, Gong, Sangkuhl, Thorn, *et al.* (2012). Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther*, 92(4), 414-417. doi:10.1038/clpt.2012.96
- Wickham. (2016). *ggplot2: Elegant Graphics for Data Analysis*: Springer-Verlag New York.
- Wiebe, Gelfand, Butler, Korbel, Fortenberry, McCabe, *et al.* (2011). Longitudinal associations of maternal depressive symptoms, maternal involvement, and diabetes management across adolescence. *J Pediatr Psychol*, 36(7), 837-846. doi:10.1093/jpepsy/jsr002
- Willer, Li, & Abecasis. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26(17), 2190-2191. doi:10.1093/bioinformatics/btq340
- Willers, Iderberg, Axelsen, Dahlström, Julin, Leksell, *et al.* (2018). Sociodemographic determinants and health outcome variation in individuals with type 1 diabetes mellitus: A register-based study. *PLoS One*, 13(6), e0199170. doi:10.1371/journal.pone.0199170
- Willi, Miller, DiMeglio, Klingensmith, Simmons, Tamborlane, *et al.* (2015). Racial-ethnic disparities in management and outcomes among children with type 1 diabetes. *Pediatrics*, 135(3), 424-434. doi:10.1542/peds.2014-1774
- Wilson. (1980). A note on the correct definition of additive deviation and dominance deviation. *Ann Hum Genet*, 44(1), 113-115. doi:10.1111/j.1469-1809.1980.tb00951.x
- Wisting, Bang, Natvig, Skriverhaug, Dahl-Jorgensen, Lask, *et al.* (2016). Metabolic Control and Illness Perceptions in Adolescents with Type 1 Diabetes. *Journal of diabetes research*, 2016, 3486094. doi:10.1155/2016/3486094
- Wojcik, Graff, Nishimura, Tao, Haessler, Gignoux, *et al.* (2019). Genetic analyses of diverse populations improves discovery for complex traits. *Nature*, 570(7762), 514-518. doi:10.1038/s41586-019-1310-4
- Wong, Dolan, Yang, & Hood. (2015). Insulin pump use and glycemic control in adolescents with type 1 diabetes: Predictors of change in method of insulin delivery across two years. *Pediatr Diabetes*, 16(8), 592-599. doi:10.1111/pedi.12221
- Wray. (2005). Allele frequencies and the  $r^2$  measure of linkage disequilibrium: impact on design and interpretation of association studies. *Twin Res Hum Genet*, 8(2), 87-94. doi:10.1375/1832427053738827
- WTCCC. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145), 661-678. doi:10.1038/nature05911
- Wu, Hilliard, Rausch, Dolan, & Hood. (2013). Family involvement with the diabetes regimen in young people: the role of adolescent depressive symptoms. *Diabet Med*, 30(5), 596-602. doi:10.1111/dme.12117
- Xiang, Yang, Chang, & Lu. (2017). The epigenetic mechanism for discordance of autoimmunity in monozygotic twins. *J Autoimmun*. doi:10.1016/j.jaut.2017.04.003

- Yamamoto, Kikuchi, Urakami, Goto, Tsubouchi, Sasaki, *et al.* (2019). Status and trends in the use of insulin analogs, insulin delivery systems and their association with glycemic control: comparison of the two consecutive recent cohorts of Japanese children and adolescents with type 1 diabetes mellitus. *Journal of pediatric endocrinology & metabolism : JPEM*, 32(1), 1-9. doi:10.1515/jpem-2018-0329
- Young, Lord, Patel, Gruhn, & Jaser. (2014). Good cop, bad cop: quality of parental involvement in type 1 diabetes management in youth. *Curr Diab Rep*, 14(11), 546. doi:10.1007/s11892-014-0546-5
- Zalloua, Azar, Delépine, Makhoul, Blanc, Sanyoura, *et al.* (2008). WFS1 mutations are frequent monogenic causes of juvenile-onset diabetes mellitus in Lebanon. *Hum Mol Genet*, 17(24), 4012-4021. doi:10.1093/hmg/ddn304
- Zarrei, MacDonald, Merico, & Scherer. (2015). A copy number variation map of the human genome. *Nat Rev Genet*, 16(3), 172-183. doi:10.1038/nrg3871
- Zduńczyk, Sendela, & Szypowska. (2014). High prevalence of depressive symptoms in well-controlled adolescents with type 1 diabetes treated with continuous subcutaneous insulin infusion. *Diabetes Metab Res Rev*, 30(4), 333-338. doi:10.1002/dmrr.2499
- Zhang, & Li. (2015). Nonparametric Risk and Nonparametric Odds in Quantitative Genetic Association Studies. *Sci Rep*, 5, 12105. doi:10.1038/srep12105
- Zimmerman, Albanese-O'Neill, & Haller. (2019). Advances in Type 1 Diabetes Technology Over the Last Decade. *Eur Endocrinol*, 15(2), 70-76. doi:10.17925/ee.2019.15.2.70
- Zins, & Goldberg. (2015). The French CONSTANCES population-based cohort: design, inclusion and follow-up. *Eur J Epidemiol*, 30(12), 1317-1328. doi:10.1007/s10654-015-0096-4
- Zuckerman-Levin, Dabaja-Younis, Ameer, Cohen, Maor, & Shehadeh. (2018). Effect of Socioeconomic Status and Ethnicity on Glycemic Control in Arab and Jewish Youth with Type 1 Diabetes Mellitus. *Rambam Maimonides medical journal*, 9(4), 1-8. doi:10.5041/RMMJ.10350

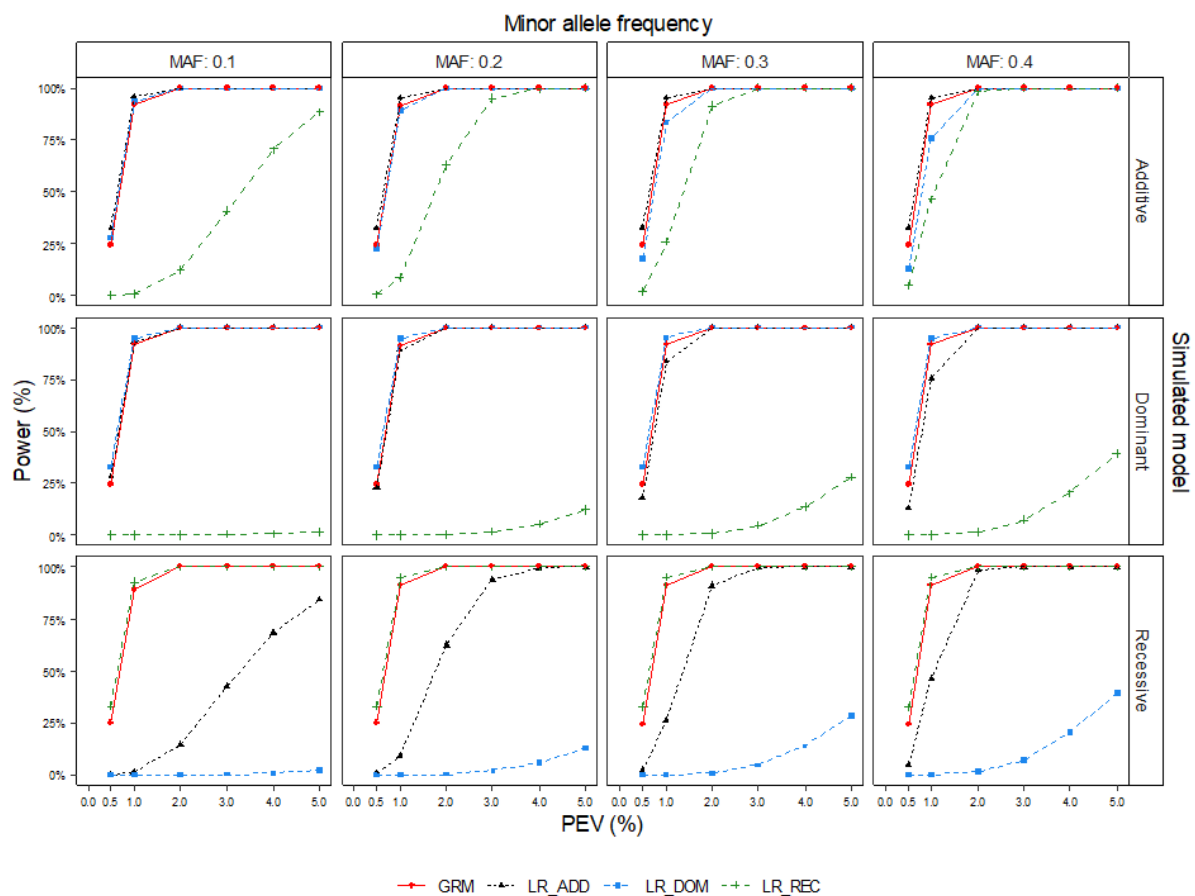
# Annexes

Annexe 1 : Données supplémentaires de l'étude des propriétés statistiques du GRM (Gloaguen *et al.*, 2019)

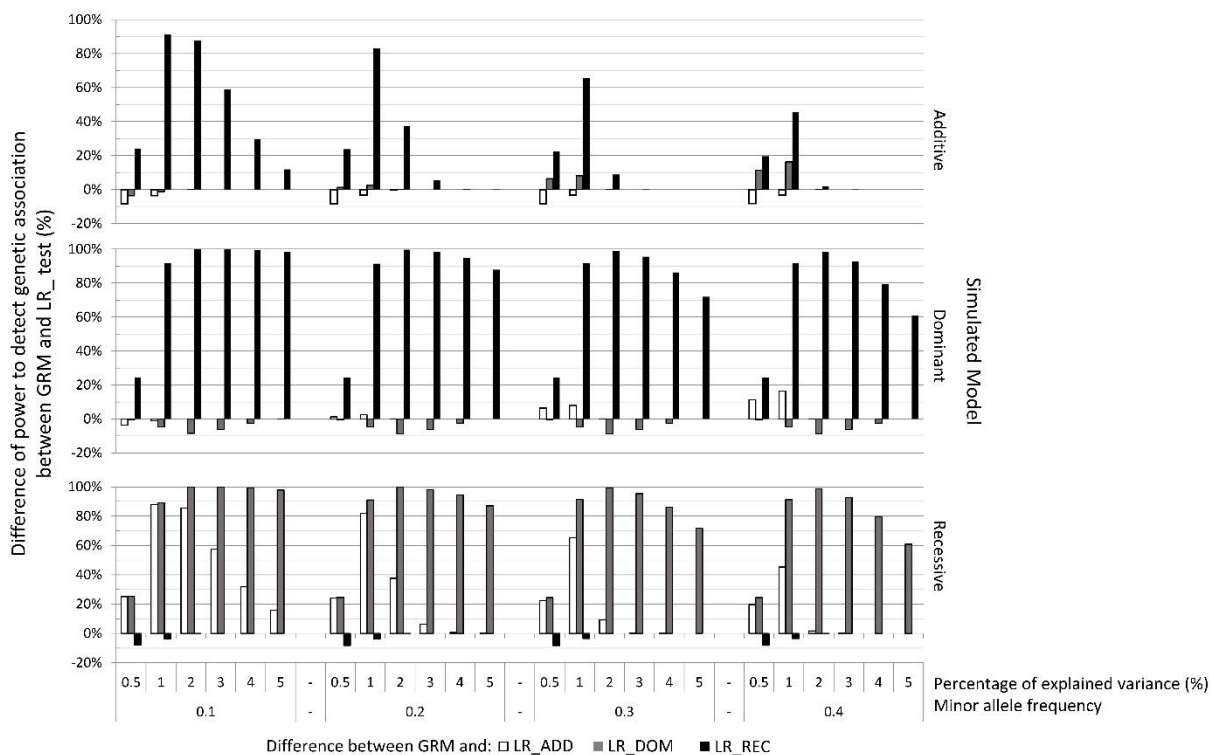
**Figure S1:** Power difference to detect association between GRM and the different LR test with a P-value threshold of 5E-8 according to the proportion of variance and minor allele frequency in a sample of 1 000 individuals.



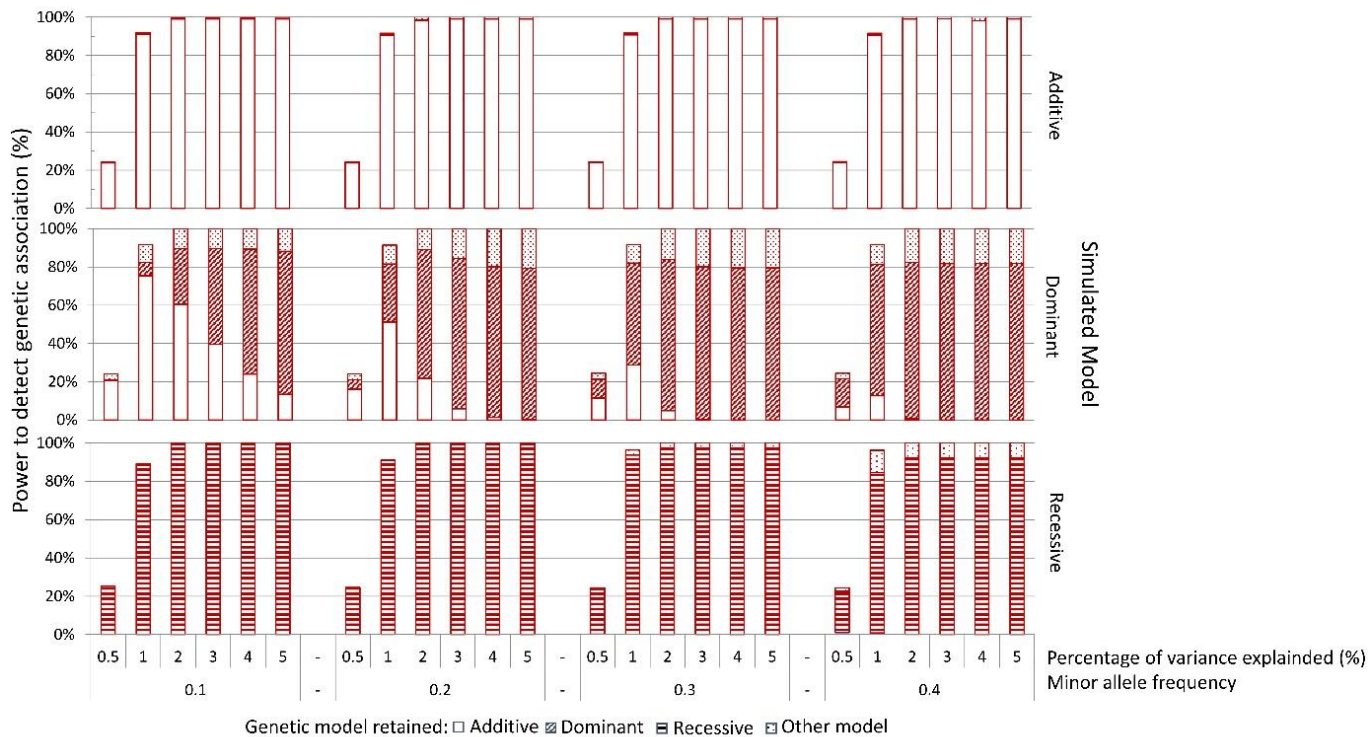
**Figure S2:** Power to detect association of GRM and LR tests, for a P-value threshold of 5E-8 for the association test, by varying mode of transmission, the minor allele frequencies and the percentage of explained variance for a sample of 5 000 individuals

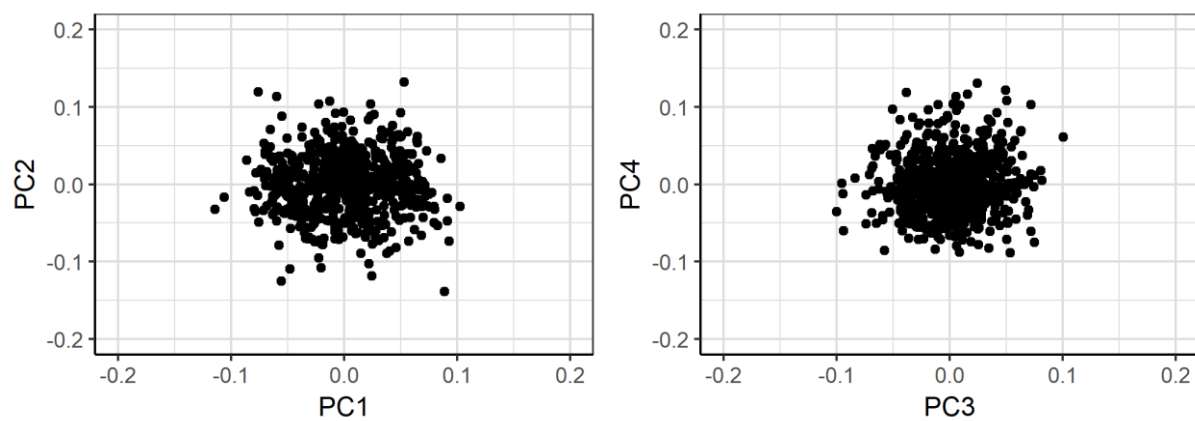


**Figure S3:** Power difference to detect association between GRM and the different LR test with a P-value threshold of 5E-8 according to proportion of variance and minor allele frequency in a sample of 5 000 individuals

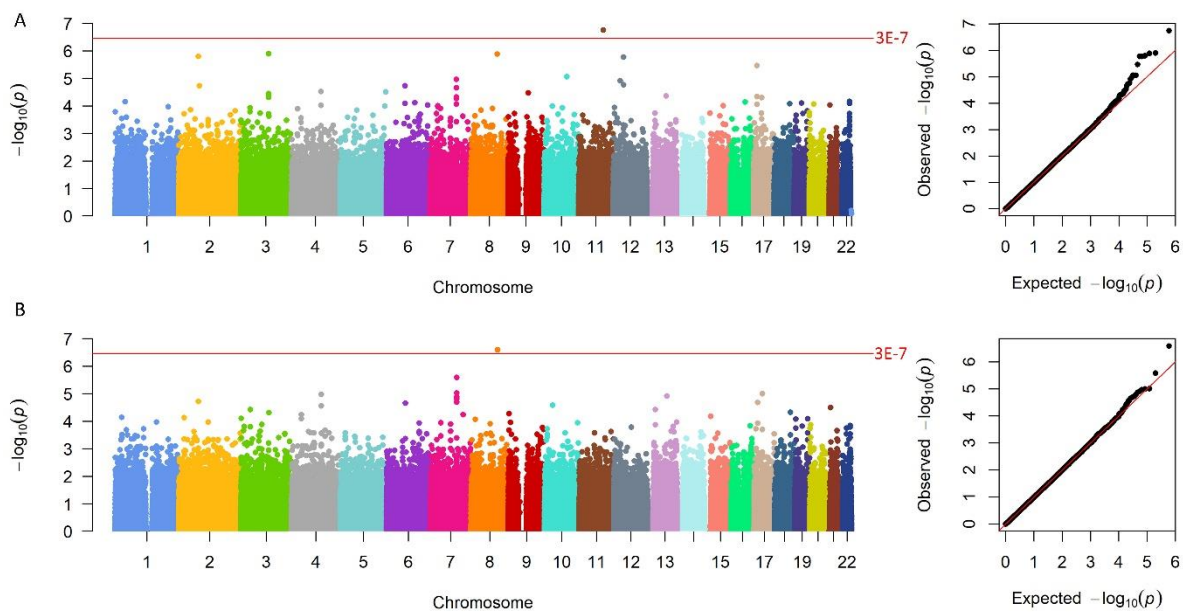


**Figure S4:** Proportion of replicates significantly associated (P-value threshold of 5E-8) retaining the additive, dominant or recessive model by using the GRM tests and using a threshold of 0.01 according to the simulated model and varying the mode of transmission, the minor allele frequency and the percentage of explained variance and for a sample of 5 000 individuals



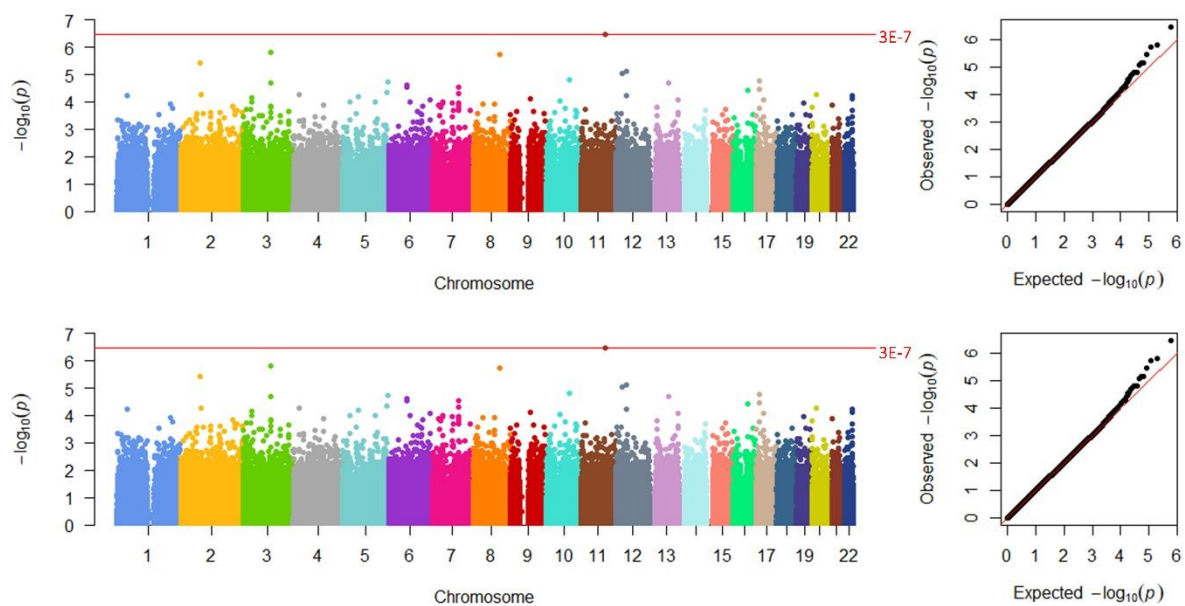
**Figure S5:** Plot of the first four principal components (PCs) of genetic data for the D.E.S.I.R. cohort

**Figure S6:** Manhattan and QQ plots of the GWAS of HbA1c on the D.E.S.I.R. cohort using A- GRM and B- LR\_ADD tests





**Figure S7:** Manhattan and QQ plots of the GWAS of HbA1c on the D.E.S.I.R. cohort taking into account for the ten first principal components using A- GRM and B- LR\_ADD tests



**Table S1:** GRM and LR tests type I error rate

Theoretical thresholds	Type I error			
	LR_ADD	LR_DOM	LR_REC	GRM
5%	0.050	0.050	0.050	0.050
1%	0.010	0.010	0.010	0.010
1.0E-5	1.0E-5	0.9E-5	1.0E-5	1.2E-5

**Table S2:** Power to detect association of GRM, MAX and allelic test, for a P-value threshold of 5E-8 using a sample of 1 000 replicates of a sample of 1 000 individuals, by varying mode of transmission, the minor allele frequencies (MAF) and the percentage of explained variance (PEV)

Model	MAF	PEV	GRM	MAX3	$\Delta$ GRM-MAX3	Allelic	$\Delta$ GRM-Allelic
Additive	1	5	93.2%	26.1%	67.1%	95.1%	-1.9%
		10	100%	83.9%	16.1%	100%	0%
		15	100%	97.4%	2.6%	100%	0%
	2	5	93.8%	80.8%	13.0%	96.1%	-2.3%
		10	100%	99.9%	0.1%	100%	0%
		15	100%	100.0%	0%	100%	0%
	3	5	94.0%	94.6%	-0.6%	95.7%	-1.7%
		10	100%	100%	0%	100%	0%
		15	100%	100%	0%	100%	0%
	4	5	92.6%	94.3%	-1.7%	95.5%	-2.9%
		10	100%	100%	0%	100%	0%
		15	100%	100%	0%	100%	0%
Dominant	1	5	93.8%	2.8%	91.0%	93.3%	0.5%
		10	100%	32.3%	67.7%	100%	0%
		15	100%	71.7%	28.3%	100%	0%
	2	5	94.2%	31.4%	62.8%	90.9%	3.3%
		10	100%	95.3%	4.7%	100%	0%
		15	100%	99.9%	0.1%	100%	0%
	3	5	94.5%	73.2%	21.3%	85.0%	9.5%
		10	100%	100%	0%	100%	0%
		15	100%	100%	0%	100%	0%
	4	5	94.0%	88.7%	5.3%	77.2%	16.8%
		10	100%	100%	0%	100%	0%
		15	100%	100%	0%	100%	0%
Recessive	1	5	82.7%	86.4%	-3.7%	3.2%	79.5%
		10	98.9%	99.2%	-0.3%	20.3%	78.6%
		15	99.9%	100%	-0.1%	43.5%	56.4%
	2	5	89.5%	92.4%	-2.9%	10.8%	78.7%
		10	100%	100%	0%	59.4%	40.6%
		15	100%	100%	0%	90.0%	10%
	3	5	92.2%	93.4%	-1.2%	25.6%	66.6%
		10	100%	100%	0%	90.8%	9.2%
		15	100%	100%	0%	99.6%	0.4%
	4	5	93.6%	95.1%	-1.5%	46.3%	47.3%
		10	100%	100%	0%	98.7%	1.3%
		15	100%	100%	0%	100%	0%

**Additional Information 1:** Script shell commands to replicate simulation analysis

(simulation of replicates, association analysis and test for models).

```
#!/bin/bash
date +%D:%k:%M:%S

#Simulation and association tests#

#Directories to store simulated data and association test results#
mkdir SIMULATION
cd SIMULATION
mkdir RESULTS

for model in add rec dom #Specify genetic model#
do for snp in 0.1 0.2 0.3 0.4 #Specify MAF#
do
    snpp=`echo $snp | sed 's/\./' `
    dirname=`echo $model"_"$snpp`
    if [ -d $dirname ] then cd $dirname else mkdir $dirname
    cd $dirname
    fi

    for var in 0 0.01 0.02 0.03 0.04 0.05 0.10 0.15 0.20 0.25 # Specify PEV#
    do
        varr=`echo $var | sed 's/\./' `
        filename=`echo $dirname"_"$varr`
        dirresults=`echo $model"_"$snpp"_"$varr"_results"`
        mkdir ../RESULTS/$dirresults
        if [ -d $dirname ] then cd $filename else mkdir $filename
        cd $filename
        fi

#Specify headers and results files names#
echo "CHR SNP BP A1 TEST NMISS BETA SE L99 U99 STAT P" >>
../../RESULTS/$dirresults/Association_GRM_$filename
echo "CHR SNP BP A1 TEST NMISS BETA SE L99 U99 STAT P" >>
../../RESULTS/$dirresults/Association_Rec_$filename
echo "CHR SNP BP A1 TEST NMISS BETA SE L99 U99 STAT P" >>
../../RESULTS/$dirresults/Association_Dom_$filename
echo "CHR SNP BP NMISS BETA SE L99 R2 T P" >> ../../RESULTS/$dirresults/Association_Add_$filename

        for i in `seq 1 100000`
        do
            echo "qtl"$i"_1 1" > Reference
            if [ "$model" = "dom" ]
            then echo "1 qtl"$i"_1 "$snp" "$snp" "$var" 1" > $filename.sim
            fi
            if [ "$model" = "add" ] then echo "1 qtl"$i"_1 "$snp" "$snp" "$var" 0" > $filename.sim
            fi

            if [ "$model" = "rec" ] then echo "1 qtl"$i"_1 "$snp" "$snp" "$var" -1" > $filename.sim
            fi

# Simulation of replicates#
plink --simulate-qt $filename.sim --seed $i --simulate-n 1000 --recode12 --out $filename"_"$i --noweb > /dev/null
```

**#Association analysis using GRM#**

```
plink --file $filename_"_$i --linear --genotypic --ci 0.99 --reference-allele Reference --out $filename --noweb > /dev/null
awk '$1!="CHR" {print $0}' $filename.assoc.linear >> ../../RESULTS/$dirresults/Association_GRM_$filename
```

**#Association analysis using GRM and LR tests#**

```
plink --file $filename_"_$i --linear --dominant --ci 0.99 --reference-allele Reference --out $filename --noweb > /dev/null
awk '$1!="CHR" {print $0}' $filename.assoc.linear >> ../../RESULTS/$dirresults/Association_Dom_$filename
```

```
plink --file $filename_"_$i --linear --recessive --ci 0.99 --reference-allele Reference --out $filename --noweb > /dev/null
awk '$1!="CHR" {print $0}' $filename.assoc.linear >> ../../RESULTS/$dirresults/Association_Rec_$filename
```

```
plink --file $filename_"_$i --assoc --qt-means --reference-allele Reference --out $filename --noweb > /dev/null
awk '$1!="CHR" {print $0}' $filename.qassoc >> ../../RESULTS/$dirresults/Association_Add_$filename
awk '$1!="CHR" {print $0}' $filename.qassoc.means >> ../../RESULTS/$dirresults/Verification_$filename
```

```
rm *qassoc *.assoc.linear *.qassoc.means *.log *.simfreq
done
date +%D:%k:%M:%S
echo "simulation & analyze "$filename" done..."
cd ..
tar --remove-files -czvf $filename.tar.gz $filename > /dev/null
echo "compression "$filename" done..."

done
cd ..
tar --remove-files -czvf $dirname.tar.gz $dirname > /dev/null
echo "compression "$dirname" done..."

done
done
```

```
date +%D:%k:%M:%S
```

**#Power estimation#**

```
cd RESULTS/
```

```
for dir in `ls -d *`
```

```
do
```

```
date +%D:%k:%M:%S
```

```
echo "calculate power"
```

```
cd $dir
```

```
#create an uniq file#
```

```
awk '$1!="CHR" {print $2, "ADD", $4, $5, $6, $7, $8, $9}' Association_Add* >> Add
```

```
awk '$1!="CHR" {print $5, $6, $7, $8, $11, $12}' Association_Rec* >> Rec
```

```
awk '$1!="CHR" {print $5, $6, $7, $8, $11, $12}' Association_Dom* >> Dom
```

```
awk '$5=="ADD" {print "MGR", $5, $6, $7, $8, $11, $12}' Association_GRM* >> GRM_ADD
```

```
awk '$5=="DOMDEV" {print $5, $6, $7, $8, $11, $12}' Association_GRM* >> GRM_DOMDEV
```

```
awk '$5=="GENO_2DF" {print $5, $6, $11, $12}' Association_GRM* >> GRM_GENO
```

```
paste Add Rec Dom GRM_ADD GRM_DOMDEV GRM_GENO >> Table
```

```
rm Add Rec Dom GRM_ADD GRM_DOMDEV GRM_GENO
```

```
echo $dir'_Associated' >> Power
```

```
echo $dir'_NA' >> NA
```

```
awk '$8!="NA" && $8<=5E-8 {print $2, $1}' Table | wc -l >> Power
```

```
awk '$8=="NA" {print $2, $1}' Table | wc -l >> NA
```

```
awk '$14!="NA" && $14<=5E-8 {print $9, $1}' Table | wc -l >> Power
```

```
awk '$14=="NA" {print $2, $1}' Table | wc -l >> NA
```

```
awk '$20!="NA" && $20<=5E-8 {print $15, $1}' Table | wc -l >> Power
```

```

awk '$2=="NA" {print $2, $1}' Table | wc -l >> NA
awk '$3!="NA" && $3<=5E-8 {print $34, $1}' Table | wc -l >> Power
awk '$3=="NA" {print $2, $1}' Table | wc -l >> NA
paste Power NA > ../Power_$$dir
rm Power NA

date +%D:%k:%M:%S
echo "calculate power "$dir" done..."

#Test for genetic model#
echo $$dir >> ../Model_$$dir
awk '$3!="NA" {print $0}' Table >> SNP_Not_NA
awk '$3=="NA" {print $0}' Table | wc -l | awk '{print $1}' >> ../Model_$$dir
awk '$3<=5E-8 {print $0}' SNP_Not_NA >> SNP_Associated
wc -l SNP_Associated | awk '{print $1}' >> ../Model_$$dir
awk '$3<=0.01' {print $0}' SNP_Associated >> SNP_other_Model
awk '$3>0.01' {print $1}' SNP_Associated | wc -l >> ../Model_$$dir
awk '$3>=($24-2.576*$25) && $3<=($24+2.576*$25) {print $1}' SNP_other_Model | wc -l >>
../Model_$$dir
awk '$3>=($24-2.576*$25) && $3<=($24+2.576*$25) {print $1}' SNP_other_Model | wc -l >>
../Model_$$dir
rm SNP_Not_NA SNP_Associated SNP_other_Model
awk '$3=="GENO" {print $0}' Verification_* > 1
awk '$3=="FREQ" {print $0}' Verification_* > 2
awk '$3=="MEAN" {print $0}' Verification_* > 3
awk '$3=="SD" {print $0}' Verification_* > 4
awk '$3=="COUNTS" {print $0}' Verification_* > 5
paste 1 2 3 4 5 >> Trait
awk '{if ($4=="2/2") {print $8, $9, $10, $11, $12, sqrt($10), $15, $16, $17, $18, $16*$10+$17*$11+$18*$12
,$21, $22, $23, $24, $28, $29, $30} else {print $8, $9, $12, $11, $10, sqrt($12), $15, $18, $17, $16,
$16*$10+$17*$11+$18*$12, $21, $24, $23, $22, $30, $29, $28}}' Trait >> Verification
rm 1 2 3 4 5
echo "calculation power model "$dir" done..."
cd ..
done
echo 'Simulation_directory_results
LR_ADD
LR_REC
LR_DOM
GRM' > PowerFile
paste PowerFile Power_* > Power
echo 'Simulation_directory_results
NA
Associated
Additive_Model
Recessive_Model
Dominant_Model' > ModelFile
paste ModelFile Model_* > GRM_Model
rm Power_* Model_* PowerFile ModelFile
date +%D:%k:%M:%S

```

**Additional Information 2:** Script shell commands to perform GRM test in a real dataset.

### # GRM analysis#

```
> GRM.sh Files File_Phenotype Asso_thresh IC_GM
```

```
# Files: Plink format files (.ped and .map files) ($1) #
# File_Phenotype: Plink format file for quantitative phenotype ($2) #
```

```
# Asso_thresh: threshold to declare significance for the association test ($3) #
# IC_GM: 1 - threshold to declare significance for genetic model tests ($4) #
# In AWK command, column are named by $i with i the number of the column #
```

### **#GRM.sh#**

```
#!/bin/bash
```

#### **# Association test in case of a test without covariates #**

```
plink --file $1 --pheno $2 --linear --genotypic --ci $4 --out $1_Assoc_GRM
awk '$5=="GENO_2DF" && $12<="3" {print $2}' $1_Assoc_GRM.assoc.linear > Assoc_$1
fgrep -wf Assoc_$1 $1_Assoc_GRM.assoc.linear > Associated_SNP_$1
uniq Assoc_$1 | wc -l
```

```
awk '$5=="ADD" {print $7, $8, $9, $10, $12}' Associated_SNP_$1 >> ADD
awk '$5=="DOMDEV" {print $7, $8, $12}' Associated_SNP_$1 >> DOMDEV
awk '$5=="GENO_2DF" {print $1, $2, $3, $4, $6, $11, $12}' Associated_SNP_$1 >> GENO
```

```
paste GENO ADD DOMDEV > GRM_$1
rm ADD DOMDEV GENO
```

#### **# Genetic Models tests #**

```
awk '$15>=(1-'$4') {print $0, "ADDITIVE"}' GRM_$1 > SNP_GM_$1
awk '$15<=(1-'$4') && $13>=$10 && $13<=$11 {print $0, "DOMINANT"}' GRM_$1 >> SNP_GM_$1
awk '$15<=(1-'$4') && $13<=$10 && $13>=$11 {print $0, "RECESSIVE"}' GRM_$1 >> SNP_GM_$1
awk '$15<=(1-'$4') && $12>(1-'$4') {print $0, "OVERDOMINANT"}' GRM_$1 >> SNP_GM_$1
awk '$8<0 && $15<=(1-'$4') && $12<=(1-'$4') && ($13<$10 || $13>-$10 || ($13<-$11 && $13>$11)) || ($8>0 && $15<=(1-'$4') && $12<=(1-'$4') && ($13<-$11 || $13>$11 || ($13<$10 && $13>-$10))' {print $0, "Other"}y'
GRM_$1 >> SNP_GM_$1
```

```
sort -n SNP_GM_$1 > SNP_GM_$1.txt
echo "CHR SNP BP A1 NMISS GRM_STAT P_GRM_2DF BETA_Add SE_BETA_Add L"$4"_BETA_Add
U"$4"_BETA_Add P_BETA_Add BETA_DomDev SE_BETA_DomDev P_BETA_DomDev Genetic_Model" >
header
```

```
cat header SNP_GM_$1.txt > Associated_SNP_GM_$1.txt
rm GRM_$1 Associated_SNP_$1 Assoc_$1 SNP_GM_$1 SNP_GM_$1.txt header
```

**# Note that the file Associated\_SNP\_GM\_\$1.txt is sorted by chromosome location. #**

#### **# Associated\_SNP\_GM\_\$1.txt reports the following items:**

```
# - CHR (Chromosome)
# - SNP (SNP identifier)
# - BP (Physical position)
# - A1 (minor allele frequency)
# - NMISS (Number of non-missing individuals included in analysis)
# - GRM_STAT (Coefficient of Likelihood ratio test statistic)
# - P_GRM_2df (p-value of the Likelihood ratio test statistic)
# - Beta_Add ( $\beta_{Add}$ )
# - SE_Beta_Add (standard error of  $\beta_{Add}$ )
# - Lxx (Lower value of the confidence interval of  $\beta_{Add}$  at xx%)
# - Uxx (Upper value of the confidence interval of  $\beta_{Add}$  at xx%)
# - P_Beta_ADD (P-value of the Wald test of  $\beta_{ADD}$ )
# - Beta_Beta_DomDev ( $\beta_{DomDev}$ )
# - SE_Beta_DomDev (standard error of  $\beta_{DomDev}$ )
# - P_Beta_DomDev (P-value of the Wald test of  $\beta_{DomDev}$ )
# - Genetic_Model (Genetic model retained by GRM)
```

Annexe 2 : Données supplémentaire de la revue systématique des facteurs influençant le taux d'HbA1c chez le jeune patients DT1 (Gloaguen *et al.*, 2018)

Cross-sectional studies					
Reference	Quality	Sub-groups of patients	Cohort size	Mean HbA1c (%)/(mmol/mol) ( $\pm$ SD if available)	Reference
Abali <i>et al</i> 2015 (Abali <i>et al.</i> , 2015) (70)	Intermediate	Glargine	85	8.9 $\pm$ 2.1 (74 $\pm$ 23.0)	Patients under glargine or detemir insulin have similar HbA1c level, but with a different insulin dose
		Detemir	32	8.5 $\pm$ 1.7 (69 $\pm$ 18.6)	
Beraki <i>et al</i> 2014 (Beraki, Magnuson, Sarnblad, Aman, & Samuelsson, 2014) (39)	High	All	4655	8.1 $\pm$ 1.2 (65 $\pm$ 13.1)	Physical activity was correlated with HbA1c level
Blackman <i>et al</i> 2014 (Blackman <i>et al.</i> , 2014) (32)	High	Pump	327	7.9 $\pm$ 0.9 (63 $\pm$ 9.8)	HbA1c level decrease after pump initiation, without an increase of severe hypoglycemia but with increased risk of diabetic ketoacidosis
		Injection	330	8.5 $\pm$ 1.1 (69 $\pm$ 12.0)	
Brancato <i>et al</i> 2014 (Brancato <i>et al.</i> , 2014) (73)	Intermediate	All	113	9.3 $\pm$ 1.8 (78 $\pm$ 19.7)	Early pump commencement provides lower and more durable HbA1c level
Campbell <i>et al</i> 2014 (Campbell <i>et al.</i> , 2014) (33)	High	HbA1c<7%	588	/ <sup>1</sup>	After adjusting for demographic and socioeconomic factors, patients with optimal HbA1c level have better diabetes management than patients with poorer glycaemic control
		HbA1c $\geq$ 9%	2684	/ <sup>1</sup>	
Cengiz <i>et al</i> 2013 (Cengiz <i>et al.</i> , 2013) (34)	High	All	13487	8.6 $\pm$ 1.5 (70 $\pm$ 16.4)	Risk of diabetic ketoacidosis and poor HbA1c level was correlated
Cherubini <i>et al</i> 2014 (Cherubini <i>et al.</i> , 2014) (49)	High	All	2025	7.9 $\pm$ 1.1 (63 $\pm$ 12.0)	No correlations were found between severe hypoglycemia and HbA1c levels, and higher risk of DKA increased HbA1c levels
Clements <i>et al</i> 2016 (Clements <i>et al.</i> , 2016) (35)	High	Ages 8–18 cohort	6574	8.8 (73)	Deterioration of HbA1c level in the 8-18 cohort through age 16
		Ages 16–26 cohort	2200	9.0 (75)	
Coulon <i>et al</i> 2016 (Coulon <i>et al.</i> , 2017) (76)	Intermediate	All	86	9.5 $\pm$ 0.2 (80 $\pm$ 2.2)	Neighborhood disadvantage was associated with poor metabolic control
de Beaufort <i>et al</i> 2013 (de Beaufort <i>et al.</i> , 2013) (50)	High	All	1133	8.0 $\pm$ 1.0 (64 $\pm$ 10.9)	Insulin regimen and management were associated with HbA1c level, but did not explain centers difference
de Bock <i>et al</i> 2012 (de Bock <i>et al.</i> , 2012) (78)	Intermediate	Non-pump	546	8.4 $\pm$ 1.1 (68 $\pm$ 12.0) <sup>2</sup>	Patients under pump therapy have lower HbA1c level
		Pump	75	8.1 $\pm$ 0.8 (65 $\pm$ 8.7) <sup>2</sup>	
Demirel <i>et al</i> 2013 (Demirel, Tepe, Esen, Buber, & Boztepe, 2013) (79)	Intermediate	All	93	8.2 $\pm$ 1.3 (66 $\pm$ 14.2)	A lot of individual and family-level factors are associated with HbA1c level



Drew <i>et al</i> 2011 (Drew <i>et al.</i> , 2011) (80)	Intermediate	All	252	8.4±1.6 (68±17.5)	HbA1c level was correlated with household income, with impact of maternal and paternal acceptance and adherence
Galler <i>et al</i> 2011 (Galler, Lindau, Ernert, Thalemann, & Raile, 2011) (82)	Intermediate	All	296	8.7 ±1.6 (72±17.5)	HbA1c level was associated with diabetes duration, socioeconomic status, and daily media consumption time, but not physical activity, on glycaemic control
Hilliard <i>et al</i> 2011 (Hilliard <i>et al.</i> , 2011) (86)	Intermediate	All	136	7.9±1.2 (63±13.1)	Paternal involvement was associated with poorer glycaemic control
Jesic <i>et al</i> 2013 (Jesic, Sajic, Bogicevic, Buljugic, & Maglajlic, 2013) (54)	High	With Microalbuminuria	16	11.0±2.0 (97±21.9)	Higher risk of microalbuminuria (MA) was associated with higher HbA1c
		Without Microalbuminuria	84	8.2±1.0 (66±10.9)	
Kalyva <i>et al</i> 2016 (Kalyva, Abdul-Rasoul, Kehl, Barkai, & Lukacs, 2016) (55)	High	Greece	84	8.3±3.6 (67±39.4)	Quality of life depends on the country
		Hungary	135	8.7±3.7 (72±40.4)	
		Kuwait	197	8.4±3.9 (68±42.6)	
Katz <i>et al</i> 2014 (Katz <i>et al.</i> , 2014) (89)	Intermediate	All	252	13.2±2.8 (121±30.6)	Improving dietary quality may improve HbA1c
Kim <i>et al</i> 2012 (Kim <i>et al.</i> , 2012) (90)	Intermediate	All	151	10.8±1.3 (95±14.2)	Patients with persistent suboptimal glycaemic control were more usually adolescents, female and with low socio-economic status.
Law <i>et al</i> 2013 (Law, Walsh, Queralt, & Nouwen, 2013) (57)	High	All	203	9.5±1.6 (80±17.5)	Parental and adolescent diabetes distress was associated with higher HbA1c level
Lukacs <i>et al</i> 2013 (Lukacs <i>et al.</i> , 2013) (91)	Intermediate	CSII	104	8.6±1.5 (70±16.4)	CSII therapy was not associated with HbA1c but with better quality of life
		MDI	135	8.8±1.6 (73±17.5)	
Mackey <i>et al</i> 2014 (Mackey <i>et al.</i> , 2014) (92)	Intermediate	All	225	8.8±1.6 (73±17.5)	More maternal depressive symptoms were associated with less parental monitoring and more conflict, and so associated with poorer glycaemic control
Mameli <i>et al</i> 2014 (Mameli <i>et al.</i> , 2014) (93)	Intermediate	All	115	8.4±1.1 (±12.0)	Using pump therapy was associated with better HbA1c level
Mianowska <i>et al</i> 2014 (Mianowska, Kaminska, Fendler, Szadkowska, & Mlynarski, 2014) (94)	Intermediate	All	224	7.5 (59)	Serum total bilirubin concentration was inversely correlated to HbA1c level
Nylander <i>et al</i> 2013 (Nylander <i>et al.</i> , 2013) (41)	Intermediate	All	175	7.2±1.1 (55±12.0)	Patients with neurodevelopmental problems have higher HbA1c level than patient without
Olsen <i>et al</i> 2015 (Olsen, Johannesen, Fredheim, & Svensson, 2015) (25)	High	All	3339	/ <sup>3</sup>	Pump treatment was associated with lower HbA1c level
Pulgaron <i>et al</i> 2014 (Pulgaron <i>et al.</i> , 2014) (97)	Intermediate	All	70	8.4±1.4 (68±15.3)	Parental numeracy was inversely correlated with HbA1c level

Redon <i>et al</i> 2014 (Redon <i>et al.</i> , 2014) (63)	High	All	7206	/ <sup>4</sup>	No insulin regimen, except premixed insulin, was associated with lower HbA1c level
Rosenbauer <i>et al</i> 2012 (Rosenbauer <i>et al.</i> , 2012) (29)	High	All	30708	8.1±1.5 (65±16.4)	Evolution of HbA1c in the last decade due to the evolution of treatment and patients education
Sands <i>et al</i> 2013 (Sands <i>et al.</i> , 2013) (99)	Intermediate	All	243	8.5±1.3 (69±14.2)	Patient and parents body mass indexes were correlated with HbA1c level
Scheuing <i>et al</i> 2015 (Scheuing <i>et al.</i> , 2015) (30)	High	German/Austrian mothers	25107	8.2±0.01 (66±0.1)	Mean HbA1c level was worse in individuals of Turkish mothers compared to patients of German or Austrian mothers.
		Turkish mothers	736	8.5±0.1 (69±1.1)	
Shalitin <i>et al</i> 2012 (Shalitin & Phillip, 2012) (100)	Intermediate	All	173	7.9±0.8 (63±8.7) <sup>5</sup>	In patients with T1D diagnosed before 6.5 years, mean HbA1c level in the first year is a predictor of future HbA1c level
Shorer <i>et al</i> 2011 (Shorer <i>et al.</i> , 2011) (101)	Intermediate	All	100	8.1±1.3 (65±14.2)	Importance of paternal involvement and authoritative non supportive parenting style during adolescence for glycaemic control
Sood <i>et al</i> 2012 (Sood <i>et al.</i> , 2012) (65)	High	All	136	7.9± 1.2 (63±13.1)	Association between poorer HbA1c level and mother-father discrepancies regarding diabetes only for youths with less involved fathers only
Stahl-Pehe <i>et al</i> 2014 (Stahl-Pehe <i>et al.</i> , 2014) (66)	High	All	840	8.3±1.4 (67±15.3)	Higher quality of life and treatment satisfaction was associated with lower HbA1c level
Telo <i>et al</i> 2015 (Telo, Volkening, Butler, & Laffel, 2015) (67)	High	All	238	8.2±1.0 (66±10.9)	Youths interested in starting continuous glucose monitoring have lower HbA1c levels than a general sample of 238 youths with type 1 diabetes
Valenzuela <i>et al</i> 2014 (Valenzuela <i>et al.</i> , 2014) (68)	High	All	439	8.5±1.5 (69±16.4)	Greater shared decision-making in caregivers of youth is associated with lower HbA1c
Wang <i>et al</i> 2011 (Wang <i>et al.</i> , 2011) (102)	Intermediate	All	225	9.1±2.0 (76±21.9)	Trajectories of HbA1c level was worse in Black population than in White and Hispanic populations in USA
Wiebe <i>et al</i> 2011 (Wiebe <i>et al.</i> , 2011) (103)	Intermediate	All	82	8.9±1.3 (74±14.2)	Mothers' depressive symptoms were associated with adherence and metabolic control in adolescence
Willi <i>et al</i> 2015 (Willi <i>et al.</i> , 2015) (37)	High	White	8841	8.4±1.4 (68±15.3)	Black patients have higher HbA1c level than White and Hispanic patients even after adjusting for socioeconomic status
		Black	697	9.6±1.9 (81±20.8)	
		Hispanic	1166	8.7±1.6 (72±17.5)	
Wisting <i>et al</i> 2016 (Wisting <i>et al.</i> , 2016) (104)	Intermediate	All	105	8.6±1.3 (70±14.2)	Illness perceptions was associated with metabolic control

**TABLE S1** Cross-sectional studies of non-genetic predictors of HbA1c levels.

1: stratification according to HbA1c level – 2: HbA1c at pump start – 3: HbA1c depending on the treatment – 4: HbA1c depending on the years of diagnosis – 5: Mean HbA1c level during puberty - 6: HbA1c level depends on the duration of T1D - 7: Depending on the age and the gender – 8: HbA1c at the end of the

study - CSII: Continuous subcutaneous insulin infusion – MDI: Multiple Daily Injection – DKA: Diabetic ketoacidosis (DKA+: with DKA and DKA-: without DKA) – NPH, glargin, detemir: type of insulin – ADIM: free-mix of aspart (A) and detemir (D) analog insulins- ANIM: premixed fixed-ratio aspart and neutral protamine Hagedorn (NPH) insulin mixture - KICK-OFF: Kids in Control of Food.

Cohort studies					
Reference	Quality	Sub-groups of patients	Cohort size	Mean HbA1c (%)/(mmol/mol) ( $\pm$ SD if available)	Reference
Bachle <i>et al</i> 2015 (Bachle <i>et al.</i> , 2015) (42)	Intermediate	All	202	8.3 $\pm$ 1.6 (67 $\pm$ 17.5)	Impact of depressive symptoms on HbA1c level depends on symptoms and gender
Bachle <i>et al</i> 2016 (Bachle, Stahl-Pehe, & Rosenbauer, 2016) (43)	High	All	819	8.3 $\pm$ 1.4 (67 $\pm$ 15.3)	Patients with eating disorder, frequent insulin restriction or both have worse glycaemic control
Baucom <i>et al</i> 2015 (Baucom <i>et al.</i> , 2015) (71)	Intermediate	All	175	8.2 $\pm$ 1.6 (66 $\pm$ 17.5)	Adolescents with depressive symptoms had poorer glycaemic control
Bizzarri <i>et al</i> 2013 (Bizzarri <i>et al.</i> , 2013) (72)	Intermediate	All	104	7.8 $\pm$ 1.0 (62 $\pm$ 10.9)	Body mass index gain between onset and end of this study was positively associated with increase of HbA1c level.
Butwicka <i>et al</i> 2013 (Butwicka, Zalepa, Fendler, Szadkowska, & Mlynarski, 2013) (75)	Intermediate	All	166	7.8 $\pm$ 1.7 (62 $\pm$ 18.6)	No impact of maternal depressive symptoms on HbA1c of patients
Chemtob <i>et al</i> 2011 (Chemtob <i>et al.</i> , 2011) (48)	High	All	193	/ <sup>6</sup>	High HbA1c level at the second year of T1D was a predictor of poor metabolic control
Deladoey <i>et al</i> 2013 (Deladoey, Henderson, & Geoffroy, 2013) (51)	High	All	1766	9.0 $\pm$ 1.2 (75 $\pm$ 13.1)	Linear association between household income and HbA1c level
Fendler <i>et al</i> 2012 (Fendler, Baranowska, Mianowska, Szadkowska, & Mlynarski, 2012) (52)	High	CSII	231	7.6 $\pm$ 1.0 <sup>2</sup> (60 $\pm$ 10.9)	Pump treatment was associated with improvement of HbA1c level
		MDI	223	8.0 $\pm$ 1.4 <sup>2</sup> (64 $\pm$ 15.3)	
Fredheim <i>et al</i> 2013 (Fredheim <i>et al.</i> , 2013) (24)	High	DKA+	2680	8.3 $\pm$ 1.3 (67 $\pm$ 14.2)	Association between diabetic ketoacidosis at onset and high long-term HbA1c level
		DKA-	284	8.5 $\pm$ 1.4 (69 $\pm$ 15.3)	
Fortenberry <i>et al</i> 2014 (Fortenberry <i>et al.</i> , 2014) (81)	Intermediate	All	213	8.4 $\pm$ 1.6 (68 $\pm$ 17.5)	Better illness perception was associated with better HbA1c level
Hanberger <i>et al</i> 2014 (Hanberger <i>et al.</i> , 2014) (40)	High	All	8020	/ <sup>7</sup>	HbA1c level at onset predicted HbA1c several years later
Helgeson <i>et al</i> 2012 (Helgeson, Becker, Escobar, & Siminerio, 2012) (84)	Intermediate	All	132	8.0 $\pm$ 1.3 (64 $\pm$ 14.2)	Family general stress increase HbA1c level
Hilliard <i>et al</i> 2013 (Hilliard <i>et al.</i> , 2013) (85)	Intermediate	All	257	8.8 $\pm$ 1.6 (73 $\pm$ 17.5)	Elevated family conflict and less parental monitoring are risk factors for poorer glycaemic control

Hilmi <i>et al</i> 2013 (Hilmi <i>et al.</i> , 2013) (53)	High	All	168	/ <sup>6</sup>	Bedouin ethnicity was associated with higher HbA1c level
Hofer <i>et al</i> 2014 (Hofer <i>et al.</i> , 2014) (27)	High	All	1146	8.3±1.8 (67±19.7)	Pre-pubertal HbA1c was a predictor of metabolic control in young adults
Hughes <i>et al</i> 2012 (Hughes, Berg, & Wiebe, 2012) (87)	Intermediate	All	137	8.6±1.8 (70±19.7)	When adolescents had both low emotional processing and low self-control, HbA1c was poorest
Jinno <i>et al</i> 2012 (Jinno <i>et al.</i> , 2012) (88)	Intermediate	All	90	8.3±1.2 (67±13.1)	After one year, no impact on HbA1c level of treatment change from NPH to detemir
King <i>et al</i> 2012 (King <i>et al.</i> , 2012) (56)	High	Moderate control	231	8.2 (66)	Prediction of patterns of deteriorating metabolic control with parental involvement and adolescents' psychosocial maturity predict
		Poor control	21	12.1 (109)	
Konrad <i>et al</i> 2015 (Konrad <i>et al.</i> , 2015) (28)	High	Insulin	57.487	8.6 (70)	Additional metformin therapy has no impact on HbA1c level
		Insulin + Metformin	525	9.0 (75)	
Marcovecchio <i>et al</i> 2014 (Marcovecchio, Heywood, Dalton, & Dunger, 2014) (58)	High	All	206	10.1±1.6 (87±17.5) <sup>8</sup>	Presence of microalbuminuria was associated with poorer HbA1c level
Nansel <i>et al</i> 2013 (Nansel, Lipsky, & Iannotti, 2013) (95)	Intermediate	All	340	8.4±1.1 (68±12.0)	Negative association between BMI and HbA1c
Redondo <i>et al</i> 2014 (Redondo <i>et al.</i> , 2014) (64)	High	All	857	7.8±1.5 (62±16.4)	Better HbA1c level after 1 year of T1D was associated with Caucasian ethnicity, higher socioeconomic status, more frequent Self- Monitoring of Blood Glucose, and a low insulin requirement
Rohan <i>et al</i> 2014 (Rohan <i>et al.</i> , 2014) (98)	Intermediate	All	239	8.2±1.4 (66±15.3)	Higher risk of lower HbA1c level is associated with higher levels of family conflict, lower frequency of blood glucose monitoring, and female gender
Waller <i>et al</i> 2013 (Waller <i>et al.</i> , 2013) (69)	High	All	142	8.6±1.4 (70±15.3)	Extreme emotional regulation score was associated with higher HbA1c level
Wong <i>et al</i> 2015 (Wong, Dolan, Yang, & Hood, 2015) (105)	Intermediate	All	150	8.8±1.9 (73±20.8)	Change in insulin delivery method was associated with change of glycaemic control.
Wu <i>et al</i> 2013 (Wu, Hilliard, Rausch, Dolan, & Hood, 2013) (106)	Intermediate	All	133	9.1±1.9 (76±20.8)	Parental involvement was associated with HbA1c level

**TABLE S2** Cohort studies of non-genetic predictors of HbA1c levels.

1: stratification according to HbA1c level – 2: HbA1c at pump start – 3: HbA1c depending on the treatment – 4: HbA1c depending on the years of diagnosis – 5: Mean HbA1c level during puberty - 6: HbA1c level depends on the duration of T1D - 7: Depending on the age and the gender – 8: HbA1c at the end of the

study - CSII: Continuous subcutaneous insulin infusion – MDI: Multiple Daily Injection – DKA: Diabetic ketoacidosis (DKA+: with DKA and DKA-: without DKA) – NPH, glargin, detemir: type of insulin – ADIM: free-mix of aspart (A) and detemir (D) analog insulins- ANIM: premixed fixed-ratio aspart and neutral protamine Hagedorn (NPH) insulin mixture - KICK-OFF: Kids in Control of Food.

Case-control and clinical trial studies					
Reference	Quality	Sub-groups of patients	Cohort size	Mean HbA1c (%)/(mmol/mol) ( $\pm$ SD if available)	Reference
Brorsson <i>et al</i> 2015 (Brorsson, Viklund, Ortqvist, & Lindholm Olinder, 2015) (74)	Intermediate	CSII	216	8.3 (67) <sup>2</sup>	Treatment with CSII resulted in an improvement in HbA1c levels up to 1 year
		MDI	215	8.3 (67) <sup>2</sup>	
Cadario <i>et al</i> 2014 (Cadario <i>et al.</i> , 2014) (47)	High	Migrants	84	8.6 (70)	Higher HbA1c level was associated with migrant status
		Italians	75	7.5 (58)	
Danne <i>et al</i> 2013 (Danne <i>et al.</i> , 2013) (77)	Intermediate	Glargine	61	8.1 $\pm$ 0.9 (67 $\pm$ 9.8)	No significant effect of the Glargine or NPH insulin on HbA1c
		NPH	64	8.3 $\pm$ 1.2 (67 $\pm$ 13.1)	
Goksen <i>et al</i> 2014 (Goksen, Atik Altinok, Ozen, Demir, & Darcan, 2014) (83)	Intermediate	Controls	52	8.1 $\pm$ 1.0 (65 $\pm$ 10.9)	Carbohydrate counting may provide good metabolic control in children and adolescents with T1D
		Carbohydrate counting group	32	8.4 $\pm$ 1.5 (68 $\pm$ 16.4)	
Libman <i>et al</i> 2015 (Libman <i>et al.</i> , 2015) (36)	High	Metformin	71	8.8 $\pm$ 0.8 (73 $\pm$ 8.7)	Addition of metformin treatment has no effect on HbA1c level
		Placebo	69	8.8 $\pm$ 0.7 (73 $\pm$ 7.7)	
Mauras <i>et al</i> 2012 (Mauras <i>et al.</i> , 2012) (59)	High	Continuous Glucose Monitoring	74	7.5 $\pm$ 1.8 (58 $\pm$ 19.7)	Continuous glucose monitoring in 4 to 9 year olds did not improve glycaemic control despite a high degree of parental satisfaction with CGM.
		Control	72	7.5 $\pm$ 1.7 (58 $\pm$ 18.6)	
Nadeau <i>et al</i> 2015 (Nadeau <i>et al.</i> , 2015) (30)	High	Metformin	34	9.2 $\pm$ 1.2 (77 $\pm$ 13.1)	Addition of metformin treatment has no significant effect on HbA1c level.
		Placebo	40	9.6 $\pm$ 1.2 (81 $\pm$ 13.1)	
Petit-Bibal <i>et al</i> 2015 (Petit-Bibal <i>et al.</i> , 2015) (61)	High	ADIM	41	8.1 $\pm$ 0.8 (65 $\pm$ 8.7)	Mean HbA1c was comparable on aspart (A) and detemir (D) insulins (ADIM) and premixed fixed-ratio aspart and neutral protamine Hagedorn (NPH) patients
		ANIM	41	8.0 $\pm$ 0.6 (64 $\pm$ 6.6)	
Price <i>et al</i> 2016 (Price <i>et al.</i> , 2016) (96)	Intermediate	All	198	9.2 (77)	No effect of a 5 day structured education course on glycaemic control at 24 months
Rabbone <i>et al</i> 2014 (Rabbone <i>et al.</i> , 2014) (62)	High	All	85	7.5 $\pm$ 0.8 (58 $\pm$ 8.7)	Carbohydrate counting helped to improve glycaemic control in children with type 1 diabetes

**TABLE S3** Case-control and clinical trial studies of non-genetic predictors of HbA1c levels.

1: stratification according to HbA1c level – 2: HbA1c at pump start – 3: HbA1c depending on the treatment – 4: HbA1c depending on the years of diagnosis – 5: Mean HbA1c level during puberty - 6: HbA1c level depends on the duration of T1D - 7: Depending on the age and the gender – 8: HbA1c at the end of the

study - CSII: Continuous subcutaneous insulin infusion – MDI: Multiple Daily Injection – DKA: Diabetic ketoacidosis (DKA+: with DKA and DKA-: without DKA) – NPH, glargin, detemir: type of insulin – ADIM: free-mix of aspart (A) and detemir (D) analog insulins- ANIM: premixed fixed-ratio aspart and neutral protamine Hagedorn (NPH) insulin mixture - KICK-OFF: Kids in Control of Food



Gene	Associations in the EBI GWAS catalog*	Biological Process (GO) **	Classification
PTPN22	<b>Alopecia areata (25608926); Antineutrophil cytoplasmic antibody-associated vasculitis (28029757); Breast cancer (23535729); Crohn's disease (18587394; 21102463; 23128233; 28067908); Gut microbiota (27694959); Hypothyroidism (22493691; 27182965); Immun cells numbers (27863252); Myasthenia gravis (23055271; 26562150); Obesity-related traits (23251661); Pediatric autoimmune diseases (26301688); Rheumatoid arthritis (17554300; 17804836; 18794853; 19503088; 20453842; 21156761; 24390342; 24449572); Systemic lupus erythematosus (19838195; 26502338; 27193031; 27399966); Thyroid peroxidase antibody levels (24586183); Type 1 diabetes and autoimmune thyroid diseases (25936594); Type 1 diabetes (17554260; 17554300; 19430480; 17632545; 18978792; 21829393); Vitiligo (20410501; 27723757)</b>	T cell receptor signaling pathway; T cell costimulation; Regulation of immune response	Autoimmune-disease ; Immunity
CTLA4	<b>Alopecia areata (20596022; 25608926); Celiac disease (20190752); Graves disease (21841780); Hypothyroidism (27182965); Inflammatory bowel disease (26192919); Myasthenia gravis (25643325); Primary biliary cholangitis (28425483); Rheumatoid arthritis (19503088; 20453842; 24390342); Subcutaneous adipose tissue (29178545); Tonsillectomy (28928442); Type 1 diabetes (18978792; 19430480; 21829393); Ulcerative colitis (26192919); Vitiligo (27723757)</b>	T cell costimulation; Fc-epsilon receptor signaling pathway; Immune response-regulating cell surface receptor signaling pathway; Positive regulation of lymphocyte activation; Immune response-regulating signaling pathway; Antigen receptor-mediated signaling pathway; Positive regulation of leukocyte proliferation; phosphatidylinositol-mediated signaling; fibroblast growth factor receptor signaling pathway; positive regulation of T cell activation; regulation of lymphocyte proliferation; cellular response to fibroblast; growth factor stimulus; epidermal growth factor receptor signaling pathway; positive regulation of lymphocyte proliferation; B cell receptor signaling pathway; neurotrophin TRK receptor signaling pathway; immune response-activating signal transduction; immune response; positive regulation of cell differentiation; regulation of immune response; cellular response to oxygen-containing compound	Autoimmune-disease ; Immunity
TLR3	Breast size (22747683)	toll-like receptor signaling pathway; toll-like receptor 3 signaling pathway; toll-like receptor 4 signaling pathway; TRIF-dependent toll-like receptor signaling pathway; positive regulation of innate immune response; innate immune response; positive regulation of I-kappaB kinase; NF-kappaB signaling; I-kappaB kinase; NF-kappaB signaling; toll-like receptor 9 signaling pathway; positive regulation of type I interferon production; cytoplasmic pattern recognition receptor signaling pathway; positive regulation of NF-kappaB transcription factor activity; toll-like receptor TLR1:TLR2 signaling pathway; toll-like receptor TLR6:TLR2 signaling pathway; toll-like receptor 2 signaling pathway; activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway; MyD88-dependent toll-like receptor signaling pathway; regulation of immune response; positive regulation of interferon-beta production; response to other	Immunity

		organism; defense response to other organism; positive regulation of cytokine production; defense response to virus; immune response; positive regulation of interleukin-6 production; positive regulation of immune system process; toll-like receptor 5 signaling pathway; toll-like receptor 10 signaling pathway; positive regulation of intracellular signal transduction; response to bacterium	
HLA	<p>Acute-on-chronic liver failure in hepatitis B (28130311); Allergy and Allergic disease (23817569; 27182965; 29083406); Allergic sensitization (23817571); <b>Alopecia areata (25608926)</b>; Alzheimer's disease (late onset) (24162737); Atopic dermatitis (26482879); <b>Arthritis (juvenile idiopathic) (18576341)</b>; Asparaginase hypersensitivity in acute lymphoblastic leukemia (25987655); Asthma (21804548); Asthma and hay fever (24388013); <b>Body mass index (28892062)</b>; <b>C-peptide levels in type I diabetes (29404672)</b>; <b>C-reactive protein levels or total cholesterol levels (27286809)</b>; <b>Celiac disease (20190752)</b>; <b>Cervical cancer (28806749)</b>; <b>Childhood ear infection (28928442)</b>; <b>Chronic hepatitis B infection (21750111; 25802187)</b>; <b>Chronic hepatitis C infection (24376798)</b>; Chronic lymphocytic leukemia (23770605; 26956414); Cold sores (28928442); Coronary heart disease (21971053); <b>Crohn's disease (23850713)</b>; Diastolic blood pressure (28135244); Drug-induced liver injury (21570397; 26959717); Emphysema imaging phenotypes (26030696); <b>Epstein-Barr virus immune response (EBNA-1) (23326239)</b>; <b>Enteric fever (25383971)</b>; Esophageal squamous cell carcinoma (25129146); Follicular lymphoma (20639881; 21533074); <b>Graves' disease (21841780)</b>; <b>HDL cholesterol (25961943)</b>; Height (20881960); <b>Hemoglobin concentration (27863252)</b>; <b>HPV seropositivity (21896673)</b>; Hypothyroidism (22493691; 27182965); <b>IgE grass sensitization (22036096)</b>; <b>IgE levels (23146381)</b>; <b>Immune cells numbers (27863252)</b>; <b>Immune response to anthrax vaccine (22658931)</b>; <b>Immunoglobulin A vasculitis (28698626)</b>; Iron status biomarkers (28334935); Knee osteoarthritis (20305777); <b>Leishmaniasis (23291585)</b>; Lung cancer in ever smokers (28604730); Lung function(28166213); Mixed cellularity Hodgkin lymphoma (29196614); Mixed cryoglobulinemia vasculitis in chronic hepatitis C infection (25030430); Mosquito bite size (28199695); <b>Multiple sclerosis (18941528)</b>; <b>Myasthenia gravis (25643325; 26562150)</b>; Neurofibrillary tangles (25188341); Nodular sclerosis Hodgkin lymphoma (29196614); Oligoclonal band status in multiple sclerosis (25616667); Oral cavity and pharyngeal cancer (27749845); Oropharynx cancer (27749845); Parental lifespan (29030599); Prostate cancer (23535732); Peanut allergy (25710614); Primary biliary cholangitis (19458352; 20639880; 21399635; 26394269); Pulmonary function (21946350; 23284291); Response to interferon beta therapy (21502966); <b>Rheumatoid arthritis (23918589; 24532677; 24782177; 28118524)</b>; <b>Scarlet fever</b></p>	regulation of T cell activation	<b>Autoimmune-disease; Immunity; Metabolism</b>

	(28928442); Schizophrenia (23212062; 26198764); Serum IgE levels (25488688); Skin and soft tissue Small cell lung carcinoma (28604730); Staphylococcus aureus infection (26450422); Systemic lupus erythematosus (19838193; 24871463; 26502338; 26606652; 26316170); Systemic sclerosis (20383147; 21750679; 21779181); Thionamide-induced agranulocytosis in Graves' disease (26151496); Tonsillectomy (28928442); Tuberculosis (29036319); Type 1 diabetes (17632545); Type 2 diabetes (27189021); Ulcerative colitis (19122664; 24837172); Waist circumference adjusted for body mass index (28448500)		
TNF	AIDS progression (19115949); Autism spectrum disorder or schizophrenia (28540026); Blood protein levels (28240269); Crohn's disease (21102463); Cervical cancer (28806749); Inflammatory bowel disease (28067908); Neonatal lupus (20662065); Psychosis (atypical) (24132900); Ulcerative colitis (28067908)	positive regulation of I-kappaB kinase/NF-kappaB signaling; positive regulation of NF-kappaB transcription factor activity; toll-like receptor 3 signaling pathway; TRIF-dependent toll-like receptor signaling pathway; death-inducing signaling complex assembly; tumor necrosis factor-mediated signaling pathway; toll-like receptor 4 signaling pathway; cellular response to tumor necrosis factor; positive regulation of defense response; NIK/NF-kappaB signaling; extrinsic apoptotic signaling pathway; cellular response to cytokine stimulus; positive regulation of immune response; apoptotic process; regulation of defense response; I-kappaB kinase/NF-kappaB signaling; immune response; cell surface receptor signaling pathway	Autoimmune-disease; Immunity
IL6	C-reactive protein (21196492); Cerebrospinal fluid clusterin levels in APOE4-carriers (26545630); Dialysis-related mortality (21546767); Fibrinogen (23969696); Longevity (26912274)	response to cytokine; response to lipopolysaccharide; cellular response to cytokine stimulus; cellular response to interleukin-6; cellular response to organic substance; interleukin-6-mediated signaling pathway; response to organonitrogen compound; response to lipid; response to bacterium; response to hormone; cellular response to hormone stimulus; response to glucocorticoid; cellular response to endogenous stimulus; response to purine-containing compound; positive regulation of defense response; cytokine-mediated signaling pathway; response to steroid hormone; response to organic cyclic compound; positive regulation of cell differentiation; regulation of immune response; defense response; response to nitrogen compound; multi-organism process; aging; regulation of cell differentiation; regulation of inflammatory response; response to organic substance; positive regulation of leukocyte differentiation; response to organophosphorus; response to alcohol; acute-phase response; positive regulation of macromolecule metabolic process; positive regulation of cytokine production; positive regulation of cellular metabolic process; regulation of innate immune response; positive regulation of smooth muscle cell proliferation; positive regulation of response to stimulus; positive regulation of cellular process; regulation of defense response; immune response; cell surface receptor signaling pathway; inflammatory response; positive regulation of cell proliferation;	Immunity

		response to peptide hormone; positive regulation of multicellular organismal process; <b>regulation of immune system process</b> ; regulation of apoptotic process; positive regulation of transcription, DNA-templated; <b>regulation of response to external stimulus</b> ; regulation of cell death	
ERBB3	<b>Alopecia areata (25608926)</b> ; Anorexia nervosa (28494655); Cognitive function (25644384); Educational attainment (25201988); Hypothyroidism (27182965); <b>Polycystic ovary syndrome (26416764)</b> ; Psoriasis (20953189); <b>Type 1 diabetes (17554260; 17554300; 18198356; 18978792; 19430480; 21829393)</b> ; <b>Vitiligo (22951725)</b> ; <b>Obesity-related traits (23251661)</b>	fibroblast growth factor receptor signaling pathway; Fc-epsilon receptor signaling pathway; cellular response to fibroblast growth factor stimulus; epidermal growth factor receptor signaling pathway; neurotrophin TRK receptor signaling pathway; phosphatidylinositol-mediated signaling; <b>innate immune response</b> ; cellular response to growth factor stimulus; wound healing; transmembrane receptor protein tyrosine kinase signaling pathway; positive regulation of protein kinase activity; activation of protein kinase activity; regulation of protein tyrosine kinase activity; regulation of protein phosphorylation; <b>defense response</b> ; blood coagulation; <b>cellular response to endogenous stimulus</b> ; positive regulation of protein tyrosine kinase activity; intracellular signal transduction; <b>insulin receptor signaling pathway</b> ; <b>immune system process</b> ; activation of transmembrane receptor protein tyrosine kinase activity; regulation of intracellular signal transduction; MAPK cascade; <b>cellular response to insulin stimulus</b>	<b>Autoimmune-disease;</b> <b>Metabolism;</b> <b>Immunity</b>
ACE	Alzheimer's disease (late onset) (24162737); Cerebrospinal fluid levels of Alzheimer's disease-related proteins (25340798); Diastolic blood pressure (28135244, 28739976); Metabolite levels (23281178, 24625756, 24816252); <b>Metabolic traits (21886157)</b>	regulation of vasodilation; regulation of blood volume by renin-angiotensin; renin-angiotensin regulation of aldosterone production; regulation of system process; regulation of systemic arterial blood pressure by hormone; regulation of blood vessel size; regulation of systemic arterial blood pressure by renin-angiotensin; regulation of renal sodium excretion; positive regulation of reactive oxygen species metabolic process; regulation of blood vessel size by renin-angiotensin; regulation of oxidoreductase activity; renal system process; positive regulation of vasodilation; blood vessel remodeling; blood circulation; regulation of NAD(P)H oxidase activity; positive regulation of oxidoreductase activity; regulation of response to wounding	Unknown

APOE	<p>Age-related macular degeneration (23326517, 23455636, 26691988); Alzheimer's disease (17474819, 17975299, 17998437, 18823527, 19125160, 19734902, 19734903, 20061627, 20885792, 20460622, 21627779, 22005931, 22245343, 22832961, 23565137, 24755620, 24770881, 23844046, 26830138); Alzheimer's disease biomarkers (23562540); Apolipoprotein Levels (23031429); <b>Blood metabolite ratios (24816252); Body fat percentage (26833246);</b> Brain imaging (20100581); Cardiovascular risk factors (20838585); C-reactive protein (18439548, 18439552, 19567438, 21196492); Cerebrospinal fluid AB1-42 levels (28247064, 28641921); Cerebrospinal T-tau levels (28247064); Cerebrospinal fluid t-tau:AB1-42 ratio (28641921); Cerebrospinal AB1-42 levels in Alzheimer's disease dementia (25027320); Cerebrospinal P-tau181p levels (28247064); Cerebrospinal fluid t-tau levels (28641921); <b>Cholesterol (17463246, 18193043, 18193044, 18802019, 19060906, 19060911, 20686565, 20864672, 21909109, 23067351, 23726366, 24097068, 25961943, 26763881, 28270201, 28334899, 28548082);</b> Cingulate cortical amyloid beta load (26421299); Coronary artery disease (26343387; 28714975, 29212778); Cognitive decline (22054870, 24468470, 28078323); Dementia with Lewy bodies (29263008); Gut microbiota (27694959); High light scatter reticulocyte percentage of red cells (27863252); High light scatter reticulocyte count (27863252); Hippocampal atrophy (22745009); Ideal cardiovascular health (27179730); Immature fraction of reticulocytes (27863252); Lewy body disease (25188341); Lipoprotein (a) levels (28512139); Lipoprotein phospholipase A2 activity in cardiovascular disease (28753643); <b>Lipid metabolism phenotypes (19936222, 22286219);</b> Lipoprotein-associated phospholipase A2 activity and mass (23118302); Lipoprotein-associated phospholipase A2 activity change in response to darapladib treatment in cardiovascular disease (28753643); <b>Lipid traits (24023260);</b> Longevity (21418511, 24688116); Metabolic syndrome (22399527); Metabolite levels (22916037, 27005778); Parental lifespan (29030599); Platelet count (27863252); Posterior cortical atrophy and Alzheimer's disease (26993346); Psychosis and Alzheimer's disease (22005930); <b>Red cell distribution width (27863252); Response to statins (LDL cholesterol change) (22331829, 25350695);</b> Resting heart rate (27798624); Reticulocyte count (27863252); Reticulocyte fraction of red cells (27863252); Sphingolipid levels (22359512; 26068415); <b>Triglyceride levels (19060911, 20686565, 25961943, 26582766, 28334899); Type 2 diabetes (27189021, 28869590);</b> Verbal memory performance (immediate recall change) (28800603); Verbal declarative memory (25648963); <b>Very long-chain saturated fatty acid levels (fatty acid 20:0) (25378659)</b></p>	<p>retinoid metabolic process; phototransduction, visible light; regulation of plasma lipoprotein particle levels; plasma lipoprotein particle remodeling; acylglycerol catabolic process; <b>lipid homeostasis; plasma lipoprotein particle assembly; cholesterol transport; plasma lipoprotein particle clearance; cholesterol homeostasis; cholesterol efflux; triglyceride catabolic process; lipid metabolic process; triglyceride-rich lipoprotein particle remodeling; regulation of cholesterol transport; triglyceride metabolic process; cellular lipid metabolic process; phospholipid efflux; reverse cholesterol transport; phospholipid transport; regulation of very-low-density lipoprotein particle remodeling;</b> receptor-mediated endocytosis; high-density lipoprotein particle assembly; <b>positive regulation of cholesterol esterification; steroid metabolic process; cholesterol metabolic process; very-low-density lipoprotein particle remodeling; lipid localization; regulation of lipid biosynthetic process; regulation of lipid metabolic process; regulation of lipid catabolic process; high-density lipoprotein particle remodeling;</b> alcohol metabolic process; response to abiotic stimulus; negative regulation of very-low-density lipoprotein particle remodeling; regulation of steroid metabolic process; regulation of Cdc42 protein signal transduction; <b>triglyceride homeostasis; positive regulation of lipid metabolic process; regulation of triglyceride metabolic process; lipoprotein metabolic process; high-density lipoprotein particle clearance; regulation of intestinal cholesterol absorption; lipoprotein catabolic process; positive regulation of lipid biosynthetic process; low-density lipoprotein particle remodeling; regulation of triglyceride catabolic process; small molecule metabolic process; negative regulation of cholesterol transport;</b> negative regulation of lipase activity; transport; ; organic substance catabolic process; regulation of lipoprotein lipase activity</p>	Metabolism
------	---	--	------------

SORCS1	White matter microstructure in first episode schizophrenia (left posterior cingulate cortex) (28924203); Abnormal white matter microstructure in first episode schizophrenia (multivariate analysis) (28924203); Systemic juvenile idiopathic arthritis (27927641); Methadone dose in opioid dependence/ (28115739); Endometriosis (23472165); Metabolite levels (Pyroglutamine) (23934736); Asparaginase hypersensitivity in acute lymphoblastic leukemia (25987655)		Unknown
GSC	Nonsyndromic cleft lip with cleft palate (28232668); Chin dimples (27182965); Nose size (27182965)	transcription from RNA polymerase II promoter; transcription initiation from RNA polymerase II promoter; tube development; negative regulation of fat cell proliferation; regulation of transcription from RNA polymerase II promoter; transcription, DNA-templated; transforming growth factor beta receptor signaling pathway; SMAD protein complex assembly;	Unknown

**TABLE S5** Phenotype and biological pathway associated with tested genes

\*Disease/Trait (PUBMEDID) with p-value < 1E-5 as reported in the EBI GWAS catalog

\*\* Biological Process (GO) with p-value < 1E-5 as reported in the STRING

**Annexe 3 :** Consentements envoyés aux patients et/ou à leur parents en fonction de leur âge au moment de l'étude



## FORMULAIRE D'INFORMATION ET DE CONSENTEMENT (Participants 6-10 ans)

**Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR)**

VERSION 1.0 DU 30 mars 2015

N° Gestionnaire	N° CPP	N° MESR	N° CNIL
C 13-70	2014/24NICB	DC-2008-72	914317

<p><b>Organisme responsable :</b> Inserm ITMO Santé Publique Pôle Recherche Clinique (PRC) 8 rue de la Croix Jarry 75013 Paris Cedex 13</p>	<p><b>Responsable de la recherche :</b> Flavie Mathieu Fonction : Chercheur (IR INSERM, HDR) Unité Inserm d'affiliation : INSERM UMR-S958</p> <p>Adresse : Faculté de Médecine Villemin 10 av de Verdun 75010 Paris Tel : 01 57 27 85 43 Fax : Email : <a href="mailto:flavie.mathieu@inserm.fr">flavie.mathieu@inserm.fr</a></p>
---	---





## **1. INFORMATION A L'ATTENTION DES PARENTS DU PARTICIPANT 6-10 ANS**

### **1. INFORMATION**

Ce document a pour but de vous fournir les informations écrites nécessaires à votre décision. Nous vous remercions de le lire attentivement. N'hésitez pas à poser des questions au médecin en charge de la recherche si vous voulez plus d'informations. Vous pouvez prendre le temps pour réfléchir à la participation de votre enfant à cette recherche, et en discuter avec votre médecin traitant et vos proches. En fin de document, si vous acceptez que votre enfant participe à l'étude, le médecin en charge de la recherche vous demandera de compléter l'emplacement qui vous est réservé en y apposant votre signature et la date de votre consentement.

### **2. CONSENTEMENT**

Si vous décidez de faire participer votre enfant, sachez que vous pourrez retirer à tout moment votre consentement à la recherche, sans que votre enfant ou vous-même n'encourriez aucune responsabilité ni aucun préjudice de ce fait. Cela ne changera en rien les rapports que vous et votre enfant avez avec votre médecin ou avec le médecin investigateur. Nous vous demanderons simplement d'en informer le médecin en charge de la recherche. Vous n'aurez pas à justifier votre décision.

Si vous décidez que votre enfant ne participe pas à la recherche, vous n'aurez pas à vous justifier et la qualité des soins dont il/elle doit bénéficier n'en sera pas modifiée.

### **3. CADRE GENERAL ET OBJECTIFS DE LA RECHERCHE**

Vous avez déjà, en tant que parents de patient atteint de diabète de type1(DT1) lors d'un précédent prélèvement sanguin, donné votre consentement pour que l'ADN de votre enfant soit conservé au Centre de Biologie et de Pathologie Est du Groupement Hospitalier Est dont dépend l'Hôpital Femme-Mère-Enfant. et étudié afin de mettre en évidence ou préciser des gènes qui prédisposent au diabète de type 1.

Dans le cadre du présent projet, nous souhaitons étudier le rôle des facteurs génétiques impliqués dans le contrôle de la dose d'insuline requise (DIR) pour obtenir un bon équilibre glycémique et étudier l'évolution de la DIR à 2 ans puis au cours de votre maladie.

Pour ce nouveau projet de recherche nous avons besoin de recueillir, à nouveau, votre consentement.

### **4. DEROULEMENT DE LA RECHERCHE**

**Aucun nouveau prélèvement de sang n'est prévu pour cette nouvelle recherche. Seul votre consentement pour cette nouvelle recherche, portant sur les données génétiques de votre enfant et les données médicales relatives à son diabète, vous est demandé.**

L'analyse des résultats portera sur environ 1000 patients et est prévue pour se dérouler sur une période de 5 ans. Par ailleurs, aucune indemnisation n'est prévue dans le cadre de cette recherche.

### **5. DROITS D'INFORMATION CONCERNANT LA SANTE DE VOTRE ENFANT**

Vous avez droit d'avoir communication en cours et à l'issue de la recherche des informations concernant la santé de votre enfant, conformément à l'article L.1122-1 du Code de la Santé Publique.

Si une ou plusieurs anomalies apparaissent dans ses résultats d'analyse, nous vous en informerons et nous vous conseillerons. Ces résultats seront transmis au médecin de votre choix.

### **6. DROIT D'ETRE INFORME DES RESULTATS GLOBAUX**

Vous avez le droit d'être informé(e) des résultats globaux de la recherche à l'issue de celle-ci, conformément au dernier alinéa de l'article L.1122-1 du Code de la Santé Publique, par le Pr Nicolino soit à l'issue des analyses.

Les résultats de cette recherche peuvent être présentés à des congrès ou dans des publications scientifiques. Cependant, les données personnelles de votre enfant ne seront aucunement identifiables car elles auront été préalablement rendues confidentielles grâce à un codage particulier qui ne mentionne ni son nom, ni son prénom.





## **7. CONFIDENTIALITE ET TRAITEMENTS DES DONNEES INFORMATISEES**

Dans le cadre de la recherche médicale à laquelle l'Inserm vous propose la participation de votre enfant, un traitement de ses données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de l'objectif de cette dernière.

A cette fin, les données médicales, le concernant, et les données relatives à vos origines ethniques collectées par le Pr Nicolino (pays de naissance des 4 grands parents de votre enfant) seront transmises de manière sécurisée à l'équipe de recherche de l'INSERM UMR-S958 située à la Faculté de Médecine Villemin 10 av de Verdun 75010 Paris. Ces données seront identifiées par un numéro de code spécifique aux analyses génétiques. Ces données pourront également, dans des conditions assurant leur confidentialité, être transmises aux Autorités de Santé Françaises.

Conformément aux dispositions de la loi relative à l'informatique, aux fichiers et aux libertés, vous disposez d'un droit d'accès et de rectification de ces données. Vous disposez également d'un droit d'opposition à la transmission des données couvertes par le secret professionnel susceptibles d'être utilisées dans le cadre de cette recherche et d'être traitées. Vous pouvez également accéder directement ou par l'intermédiaire d'un médecin de votre choix, à l'ensemble des données médicales de votre enfant en application des dispositions de l'article L. 1111-7 du Code de la Santé Publique. Ces droits s'exercent auprès du médecin qui suit votre enfant dans le cadre de la recherche et qui connaît son identité.

## **8. INFORMATIONS SUR LE DEVENIR DE L'ADN DE VOTRE ENFANT A L'ISSUE DE LA RECHERCHE**

### ***En fin de recherche***

Après utilisation d'une fraction d'ADN pour les besoins de cette recherche, l'ADN restant sera conservé dans la banque d'ADN, sous la responsabilité du Pr Nicolino. La banque originale d'ADN est hébergée au Centre de Biologie et de Pathologie Est du Groupement Hospitalier Est dont dépend l'Hôpital Femme-Mère-Enfant. Une déclaration de collection a été faite par ce centre sous la référence DC-2011-1437.

En cas de nouvelle recherche programmée, les données génétiques de votre enfant ne pourront en aucun cas être étudiées sans un nouveau recueil de consentement écrit de votre part.

D'autre part, vous pouvez librement, et à tout moment, sans conséquence pour la prise en charge médicale de votre enfant, vous opposer à l'utilisation ultérieure de ses données génétiques à but de recherche, en vous adressant au médecin qui le suit dans le cadre de cette recherche.

## **9. CADRE LEGISLATIF DE LA RECHERCHE MEDICALE**

Cette recherche est réalisée conformément aux dispositions légales en vigueur régissant les recherches médicales. Ainsi cette recherche est réalisée conformément aux dispositions :

- De la loi relative à la bioéthique (loi du 7 juillet 2011 modifiée par la loi du 6 août 2013)
- De la loi relative l'Informatique, aux fichiers et aux libertés (loi du 6 janvier 1978 modifiée par la loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel du 6 août 2004)
- Du Code de la Santé Publique

Cette recherche a reçu l'avis favorable du Comité de Protection des Personnes Ile de France IV le 24 avril 2014 et l'autorisation de la CNIL le 16 février 2015.

Pour pouvoir participer à cette recherche médicale votre enfant doit être affilié(e) à un régime de Sécurité Sociale, à la Couverture Médicale Universelle (CMU) ou à tout régime équivalent.



## **2. INFORMATION A L'ATTENTION DU PARTICIPANT 6-10 ANS**

Nous te proposons de participer à une étude intitulée : **Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR).**

Tu es libre d'accepter ou de refuser. Avant de le signer (avec tes parents) il faut que tu aies bien lu et compris ce qui suit. Tu peux demander au docteur de t'expliquer tout ce que tu ne comprends pas.

Le diabète de type 1 est une maladie compliquée. Avec ta famille, tu as participé il y a quelques années à une étude des gènes du diabète. En effet, certains caractères physiques, morphologiques ou même moins visibles sont héréditaires. Ils sont transmis des parents aux enfants par ce que l'on nomme les gènes. Nous cherchons ces gènes qui pourraient expliquer ta maladie.

Aujourd'hui nous te proposons de participer à la suite de cette recherche, qui porte sur la réponse au traitement par l'insuline (hormone qui contrôle taux de sucre dans le sang (appelé glycémie)). Certains enfants ont besoin de faibles doses d'insuline pour équilibrer leur glycémie. D'autres enfants ont une glycémie qui est difficile à équilibrer. Pour comprendre, nous avons besoin de regarder les doses d'insuline et l'équilibre glycémique chez un grand nombre d'enfants diabétiques au cours de leur vie.

Si tu es d'accord pour participer à cette recherche, nous utiliserons les informations de ton dossier médical et tes gènes. Si besoin, nous ferons d'autres études génétiques sur ton ADN (longue molécule qui contient l'information génétique d'un individu) stocké à l'hôpital.

**Dans cette recherche il n'y a pas besoin de faire de prise de sang.**

Bien sûr, c'est toi qui choisis de participer ou non à cette étude. Si tu acceptes, tu devras signer un document avec tes parents. Tu as aussi le droit de changer d'avis après avoir donné ton accord. Dans ce cas, dis le à tes parents et au docteur. Tout ce qui te concerne sera retiré de l'étude.

*Merci, ton aide est très importante pour nous !*

*Pr. Marc Nicolino  
Dr. Flavie Mathieu*



### **3. RECUEIL DU CONSENTEMENT DES TITULAIRES DE L'AUTORITE PARENTALE**

Je soussigné(e) \_\_\_\_\_ atteste avoir bien lu et pris connaissance des informations relatives à la participation de mon enfant à la recherche intitulée « **Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise (DIR) pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR)** » exposées par écrit sur les pages précédentes (p 2 à 4) et transmises oralement par le Pr Nicolino en charge de la recherche, et avoir obtenu toutes les réponses aux questions que je lui ai posées.

- J'ai bien été informé(e) que la participation de mon enfant à cette recherche consiste à donner mon consentement pour étudier ses données génétiques relatives au diabète corrélées à l'évolution de la DIR.
- J'ai été avisé(e) qu'aucune indemnisation n'est prévue pour cette recherche.
- J'ai bien noté que le droit d'accès aux données de mon enfant, prévu par la loi du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés, modifiée par la loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel du 6 août 2004 s'exerce à tout moment auprès du médecin qui suit mon enfant dans le cadre de la recherche et qui connaît son identité. Je pourrai exercer mon droit de rectification et d'opposition auprès de ce même médecin.
- J'ai eu suffisamment de temps pour réfléchir à la participation de mon enfant à cette recherche.
- Je suis parfaitement conscient(e) que je peux retirer à tout moment mon consentement de participation à cette recherche quelque soient mes raisons et sans avoir à m'en justifier, sans que mon enfant n'en supporte aucune responsabilité ni préjudice. J'en informerai simplement le médecin en charge de la recherche.

#### **CONSENTEMENT RELATIF AUX DONNEES PERSONNELLES**

J'accepte que les données concernant mon enfant, enregistrées à l'occasion de cette recherche, comportant des données médicales, génétiques et ethniques puissent faire l'objet d'un traitement informatique par le gestionnaire ou pour son compte. J'accepte également que l'ensemble de mon dossier médical puisse être consulté par les personnes habilitées dans le cadre de cette recherche, dans le respect de la confidentialité des données de mon enfant et de son identité.

#### **CONSENTEMENT RELATIF A L'ADN**

J'ai été informé(e) que l'ADN de mon enfant reste sous la responsabilité du Pr. Nicolino après la fin de cette recherche. J'accepte par ma signature que les analyses génétiques soient effectuées sur l'ADN de mon enfant si elles n'ont pas encore été effectuées afin de répondre à l'objectif de cette recherche. Je serais éventuellement sollicité(e) pour signer un nouveau consentement dans le cadre d'une nouvelle recherche portant sur les données génétiques liées au diabète de mon enfant. Ces recherches seront menées par l'équipe de recherche de l'INSERM UMR-S958 située à la Faculté de Médecine Villemin, 10 av de Verdun 75010 Paris.

**J'accepte librement et volontairement la participation de mon enfant à la recherche décrite ci-dessus.** Mon consentement ne décharge en rien le médecin en charge de la recherche et le gestionnaire de l'ensemble de leurs responsabilités et je conserve tous mes droits garantis par la loi.

#### **Personnes donnant leurs consentements :**

**Père :** NOM, Prénom :

**Mère :** NOM, Prénom :

Date :

Date :

Signature :

Signature :

**Signature de l'enfant :**

**Médecin en charge de la recherche :** NOM, Prénom :

☎ :

Date :

Signature :

La dernière page de ce document doit comporter les signatures de l'investigateur et des titulaires de l'autorité parentale sollicités et doit être datée de la main des titulaires de l'autorité parentale qui ont consenti dans l'emplacement prévu et qui leur est réservé.

Ce document d'information et de consentement est établi en trois exemplaires originaux : un exemplaire est remis à chacun des titulaires de l'autorité parentale, et un exemplaire est conservé pendant la durée légale de conservation des documents de la recherche, par le médecin en charge de la recherche.



## FORMULAIRE D'INFORMATION ET DE CONSENTEMENT RECHERCHE NON INTERVENTIONNELLE 11-17 ans

Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR)

VERSION V2.0 DU 30 mars 2015

N° Gestionnaire	N° CPP	N° MESR	N°CNIL
C 13-70	2014/24NICB	DC-2008-72	914317

<p><b>Organisme responsable :</b> Inserm ITMO Santé Publique Pôle Recherche Clinique (PRC) 8 rue de la Croix Jarry 75 013 Paris 13</p>	<p><b>Responsable de la recherche :</b> Dr Flavie Mathieu Fonction : Chercheur (IR INSERM, HDR) Unité Inserm d'affiliation : INSERM UMR-S958</p> <p>Adresse : Faculté de Médecine Villemin 10 av de Verdun 75010 Paris Tel : 01 57 27 85 43 Fax : Email : <a href="mailto:flavie.mathieu@inserm.fr">flavie.mathieu@inserm.fr</a></p>
--	--



## **1. INFORMATION A L'ATTENTION DES PARENTS DU PARTICIPANT 11-17 ANS**

### **1. INFORMATION**

Ce document a pour but de vous fournir les informations écrites nécessaires à votre décision. Nous vous remercions de le lire attentivement. N'hésitez pas à poser des questions au médecin en charge de la recherche si vous voulez plus d'informations. Vous pouvez prendre le temps pour réfléchir à la participation de votre enfant à cette recherche, et en discuter avec votre médecin traitant et vos proches. En fin de document, si vous acceptez que votre enfant participe à l'étude, le médecin en charge de la recherche vous demandera de compléter l'emplacement qui vous est réservé en y apposant votre signature et la date de votre consentement.

### **2. CONSENTEMENT**

Si vous décidez de faire participer votre enfant, sachez que vous pourrez retirer à tout moment votre consentement à la recherche, sans que votre enfant ou vous-même n'encourriez aucune responsabilité ni aucun préjudice de ce fait. Cela ne changera en rien les rapports que vous et votre enfant avez avec votre médecin ou avec le médecin investigateur. Nous vous demanderons simplement d'en informer le médecin en charge de la recherche. Vous n'aurez pas à justifier votre décision.

Si vous décidez que votre enfant ne participe pas à la recherche, vous n'aurez pas à vous justifier et la qualité des soins dont il/elle doit bénéficier n'en sera pas modifiée.

### **3. CADRE GENERAL ET OBJECTIFS DE LA RECHERCHE**

Vous avez déjà, en tant que parents de patient atteint de diabète de type1(DT1) lors d'un précédent prélèvement sanguin, donné votre consentement pour que l'ADN de votre enfant soit conservé dans une banque d'ADN au sein de la collection de la biotèque des Hospices Civils de Lyon et étudié afin de mettre en évidence ou préciser des gènes qui prédisposent au diabète de type 1.

Dans le cadre du présent projet, nous souhaitons étudier le rôle des facteurs génétiques impliqués dans le contrôle de la dose d'insuline requise (DIR) pour obtenir un bon équilibre glycémique et étudier l'évolution de la DIR à 2 ans puis au cours de votre maladie.

Pour ce nouveau projet de recherche nous avons besoin de recueillir, à nouveau, votre consentement.

### **4. DEROULEMENT DE LA RECHERCHE**

**Aucun nouveau prélèvement de sang n'est prévu pour cette nouvelle recherche. Seul votre consentement pour cette nouvelle recherche, portant sur les données génétiques de votre enfant et les données médicales relatives à son diabète, vous est demandé.**

L'analyse des résultats portera sur environ 1000 patients et est prévue pour se dérouler sur une période de 5 ans.

Par ailleurs, aucune indemnisation n'est prévue dans le cadre de cette recherche.

### **5. DROITS D'INFORMATION CONCERNANT LA SANTE DE VOTRE ENFANT**

Vous avez droit d'avoir communication en cours et à l'issue de la recherche des informations concernant la santé de votre enfant, conformément à l'article L.1122-1 du Code de la Santé Publique.

Si une ou plusieurs anomalies apparaissent dans ses résultats d'analyse, nous vous en informerons et nous vous conseillerons. Ces résultats seront transmis au médecin de votre choix.

### **6. DROIT D'ETRE INFORME DES RESULTATS GLOBAUX**

Vous avez le droit d'être informé(e) des résultats globaux de la recherche à l'issue de celle-ci, conformément au dernier alinéa de l'article L.1122-1 du Code de la Santé Publique, par le Pr Nicolino soit à l'issue des analyses.

Les résultats de cette recherche peuvent être présentés à des congrès ou dans des publications scientifiques. Cependant, les données personnelles de votre enfant ne seront aucunement identifiables car elles auront été préalablement rendues confidentielles grâce à un codage particulier qui ne mentionne ni son nom, ni son prénom.



## **7. CONFIDENTIALITE ET TRAITEMENTS DES DONNEES INFORMATISEES**

Dans le cadre de la recherche médicale à laquelle l'Inserm vous propose la participation de votre enfant, un traitement de ses données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de l'objectif de cette dernière.

A cette fin, les données médicales, le concernant, et les données relatives à vos origines ethniques collectées par le Pr Nicolino (pays de naissance des 4 grands parents de votre enfant) seront transmises de manière sécurisée à l'équipe de recherche de l'INSERM UMR-S958 située à la Faculté de Médecine Villemin 10 av de Verdun 75010 Paris. Ces données seront identifiées par un numéro de code spécifique aux analyses génétiques. Ces données pourront également, dans des conditions assurant leur confidentialité, être transmises aux Autorités de Santé Françaises.

Conformément aux dispositions de la loi relative à l'informatique, aux fichiers et aux libertés, vous disposez d'un droit d'accès et de rectification de ces données. Vous disposez également d'un droit d'opposition à la transmission des données couvertes par le secret professionnel susceptibles d'être utilisées dans le cadre de cette recherche et d'être traitées. Vous pouvez également accéder directement ou par l'intermédiaire d'un médecin de votre choix, à l'ensemble des données médicales de votre enfant en application des dispositions de l'article L. 1111-7 du Code de la Santé Publique. Ces droits s'exercent auprès du médecin qui suit votre enfant dans le cadre de la recherche et qui connaît son identité.

## **8. INFORMATIONS SUR LE DEVENIR DE L'ADN DE VOTRE ENFANT A L'ISSUE DE LA RECHERCHE**

### ***En fin de recherche***

Après utilisation d'une fraction d'ADN pour les besoins de cette recherche, l'ADN restant sera conservé dans la banque d'ADN, sous la responsabilité du Pr Nicolino. La banque originale d'ADN est hébergée au Centre de Biologie et de Pathologie Est du Groupement Hospitalier Est dont dépend l'Hôpital Femme-Mère-Enfant. Une déclaration de collection a été faite par ce centre sous la référence DC-2011-1437.

En cas de nouvelle recherche programmée, les données génétiques de votre enfant ne pourront en aucun cas être étudiées sans un nouveau recueil de consentement écrit de votre part.

D'autre part, vous pouvez librement, et à tout moment, sans conséquence pour la prise en charge médicale de votre enfant, vous opposer à l'utilisation ultérieure de ses données génétiques à but de recherche, en vous adressant au médecin qui le suit dans le cadre de cette recherche.

## **9. CADRE LEGISLATIF DE LA RECHERCHE MEDICALE**

Cette recherche est réalisée conformément aux dispositions légales en vigueur régissant les recherches médicales. Ainsi cette recherche est réalisée conformément aux dispositions :

- De la loi relative à la bioéthique (loi du 7 juillet 2011 modifiée par la loi du 6 août 2013)
- De la loi relative l'Informatique, aux fichiers et aux libertés (loi du 6 janvier 1978 modifiée par la loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel du 6 août 2004)
- Du Code de la Santé Publique

Cette recherche a reçu l'avis favorable du Comité de Protection des Personnes Ile de France IV le 24 avril 2014 et l'autorisation de la CNIL le 16 février 2015.

Pour pouvoir participer à cette recherche médicale votre enfant doit être affilié(e) à un régime de Sécurité Sociale, à la Couverture Médicale Universelle (CMU) ou à tout régime équivalent.





## **2. INFORMATION A L'ATTENTION DU PARTICIPANT 11-17 ANS**

L'étude que nous te proposons porte sur l'étude de la réponse au traitement. Il faut que tu aies bien lu et compris ce qui suit. Tu peux poser des questions si tu le souhaites.

### **1- Cadre général et objectif de la recherche**

Nous te proposons de participer à une étude intitulée : Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR).

Tu es libre d'accepter ou de refuser. Avant de le signer (avec tes parents) il faut que tu aies bien lu et compris ce qui suit. Tu peux demander au médecin de t'expliquer tout ce que tu ne comprends pas.

Le diabète de type 1 est une maladie compliquée. Avec ta famille, tu as participé il y a quelques années à une étude des gènes du diabète. En effet, certains caractères physiques, morphologiques ou même moins visibles sont héréditaires. Ils sont transmis des parents aux enfants par ce que l'on nomme les gènes. Nous cherchons ces gènes qui pourraient expliquer ta maladie.

Aujourd'hui nous te proposons de participer à la suite de cette recherche, qui porte sur la réponse au traitement par l'insuline (hormone qui contrôle taux de sucre dans le sang (appelé glycémie)). Certains enfants ont besoin de faibles doses d'insuline pour équilibrer leur glycémie. D'autres enfants ont une glycémie qui est difficile à équilibrer. Pour comprendre, nous avons besoin de regarder les doses d'insuline et l'équilibre glycémique chez un grand nombre d'enfants diabétiques au cours de leur vie.

### **2- Déroulement de la recherche**

Si tu es d'accord pour participer à cette recherche, nous utiliserons les informations de ton dossier médical et tes données génétiques. Si besoin, nous ferons d'autres études génétiques sur ton ADN (longue molécule qui contient l'information génétique d'un individu) stocké à l'hôpital.

### **3- Risques et contraintes particulières**

**Dans cette recherche il n'y a pas besoin de faire de prise de sang.**

La participation à cette recherche ne comporte aucun risque et aucune contrainte particulière.

### **4- Indemnisation**

Aucune indemnité ou avantage financier ne sont prévus pour la participation à cette recherche.

### **5- Cadre législatif**

Ainsi cette recherche est réalisée conformément à la réglementation exigée en recherche clinique.

Cette recherche a reçu l'avis favorable du Comité de Protection des Personnes Ile de France IV le 24 avril 2014 et l'autorisation de la CNIL le 16 février 2015.

### **6- Données informatisées**

Les données recueillies au cours de cette étude demeureront strictement confidentielles. Elles ne pourront être consultées que par l'équipe médicale, les personnes dûment habilitées.

Bien évidemment tu es libre d'accepter ou de refuser de participer à cette étude. Si tu acceptes, tu devras signer un formulaire de consentement avec tes parents. Si tu deviens majeur au cours de l'étude, nous te demanderons à nouveau de signer un formulaire de consentement à ta majorité. Tu peux retirer à tout moment ton consentement de participation à cette recherche et cela quelles que soient tes raisons. Dans le cas où tu retirerais ton consentement, toutes les données te concernant seront détruites (y compris ton prélèvement de sang). Le fait de ne plus participer à cette recherche ne portera pas atteinte à tes relations avec le médecin qui te suit.

Tu peux à tout moment demander des informations complémentaires à tes parents qui se renseigneront si besoin auprès du médecin de l'étude.

Merci, ton aide est très importante pour nous !

Pr. Marc Nicolino (Chef du service d'Endocrinologie Pédiatrique de l'hôpital Femme Mère-Enfant)



### **3. RECUEIL DU CONSENTEMENT DES TITULAIRES DE L'AUTORITE PARENTALE**

Je soussigné(e) \_\_\_\_\_ atteste avoir bien lu et pris connaissance des informations relatives à la participation de mon enfant à la recherche intitulée « **Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise (DIR) pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type 1 (GENDIR)** » exposées par écrit sur les pages précédentes (p 2 à 4) et transmises oralement par le Pr Nicolino en charge de la recherche, et avoir obtenu toutes les réponses aux questions que je lui ai posées.

- J'ai bien été informé(e) que la participation de mon enfant à cette recherche consiste à donner mon consentement pour étudier ses données génétiques relatives au diabète corrélées à l'évolution de la DIR.
- J'ai été avisé(e) qu'aucune indemnisation n'est prévue pour cette recherche.
- J'ai bien noté que le droit d'accès aux données de mon enfant, prévu par la loi du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés, modifiée par la loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel du 6 août 2004 s'exerce à tout moment auprès du médecin qui suit mon enfant dans le cadre de la recherche et qui connaît son identité. Je pourrai exercer mon droit de rectification et d'opposition auprès de ce même médecin.
- J'ai eu suffisamment de temps pour réfléchir à la participation de mon enfant à cette recherche.
- Je suis parfaitement conscient(e) que je peux retirer à tout moment mon consentement de participation à cette recherche quelque soient mes raisons et sans avoir à m'en justifier, sans que mon enfant n'en supporte aucune responsabilité ni préjudice. J'en informerai simplement le médecin en charge de la recherche.

#### **CONSENTEMENT RELATIF AUX DONNEES PERSONNELLES**

J'accepte que les données concernant mon enfant, enregistrées à l'occasion de cette recherche, comportant des données médicales, génétiques et ethniques puissent faire l'objet d'un traitement informatique par le gestionnaire ou pour son compte. J'accepte également que l'ensemble de mon dossier médical puisse être consulté par les personnes habilitées dans le cadre de cette recherche, dans le respect de la confidentialité des données de mon enfant et de son identité.

#### **CONSENTEMENT RELATIF A L'ADN**

J'ai été informé(e) que l'ADN de mon enfant reste sous la responsabilité du Pr. Nicolino après la fin de cette recherche. J'accepte par ma signature que les analyses génétiques soient effectuées sur l'ADN de mon enfant si elles n'ont pas encore été effectuées afin de répondre à l'objectif de cette recherche. Je serais éventuellement sollicité(e) pour signer un nouveau consentement dans le cadre d'une nouvelle recherche portant sur les données génétiques liées au diabète de mon enfant. Ces recherches seront menées par l'équipe de recherche de l'INSERM UMR-S958 située à la Faculté de Médecine Villemin, 10 av de Verdun 75010 Paris.

**J'accepte librement et volontairement la participation de mon enfant à la recherche décrite ci-dessus.**

Mon consentement ne décharge en rien le médecin en charge de la recherche et le gestionnaire de l'ensemble de leurs responsabilités et je conserve tous mes droits garantis par la loi.

<b>Personnes donnant leurs consentements :</b>	
<b>Père</b> : NOM, Prénom :	<b>Mère</b> : NOM, Prénom :
Date :	Date :
Signature :	Signature :
<b>Signature de l'enfant :</b>	
<b>Médecin en charge de la recherche</b> : NOM, Prénom : <span style="float: right;">☎ :</span>	
Date :	
Signature :	

La dernière page de ce document doit comporter les signatures de l'investigateur et des titulaires de l'autorité parentale sollicités et doit être datée de la main des titulaires de l'autorité parentale qui ont consenti dans l'emplacement prévu et qui leur est réservé.

Ce document d'information et de consentement est établi en trois exemplaires originaux : **un exemplaire est remis à chacun des titulaires de l'autorité parentale**, et un exemplaire est conservé pendant la durée légale de conservation des documents de la recherche, par le médecin en charge de la recherche.





Institut national  
de la santé et de la recherche médicale

## FORMULAIRE D'INFORMATION ET DE CONSENTEMENT Adultes

Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR)

VERSION 1.0 DU 12 mars 2014

N° Gestionnaire	N° CPP/IRB	N° MESR	N°CNIL
C 13-70	2014/24NICB	DC-2011-1437	914317

<p><b>Organisme responsable :</b> Inserm ITMO Santé Publique Pôle Recherche Clinique (PRC) 8 rue de la Croix Jarry 75 013 Paris Cedex 13</p>	<p><b>Responsable de la recherche :</b> Flavie Mathieu Fonction : Chercheur (IR INSERM, HDR) Unité Inserm d'affiliation : INSERM UMR-S958</p> <p>Adresse : Faculté de Médecine Villemin 10 av de Verdun 75010 Paris Tel : 01 57 27 85 43 Fax : Email : <a href="mailto:flavie.mathieu@inserm.fr">flavie.mathieu@inserm.fr</a></p>
--	---



## **1. INFORMATION A L'ATTENTION DU PARTICIPANT ADULTE**

Madame, Monsieur,

Nous vous proposons de participer à la recherche intitulée : **Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR).**

L'Institut National de la Santé et de la Recherche Médicale (Inserm) est gestionnaire de cette recherche.

Le Pr. Nicolino (Chef du Service d'endocrinologie, Diabétologie et Métabolisme Pédiatriques Division de Pédiatrie de l'Hôpital Femme-Mère-Enfant des Hôpitaux de Lyon - ☎ 04 72 12 95 25) en charge de la recherche médicale est à votre disposition pour vous présenter la recherche et la façon dont vous pouvez y participer, pour répondre à toutes vos questions et pour vous expliquer ce que vous ne comprenez pas.

### **1. INFORMATION**

Ce document a pour but de vous fournir les informations écrites nécessaires à votre décision. Nous vous remercions de le lire attentivement. N'hésitez pas à poser des questions au médecin en charge de la recherche si vous voulez plus d'informations. Vous pouvez prendre le temps pour réfléchir à votre participation à cette recherche, et en discuter avec votre médecin traitant et vos proches. En fin de document, si vous acceptez de participer à cette étude, le médecin en charge de la recherche vous demandera de compléter l'emplacement qui vous est réservé en y apposant votre signature et la date de votre consentement.

### **2. CONSENTEMENT**

Votre participation est volontaire : vous êtes libre d'accepter ou de refuser de participer à cette recherche médicale.

Si vous décidez de participer, sachez que vous pourrez retirer à tout moment votre consentement à la recherche, sans encourir aucune responsabilité ni aucun préjudice de ce fait. Cela ne changera en rien les rapports que vous avez avec votre médecin ou avec le médecin investigateur. Nous vous demanderons simplement d'en informer le médecin en charge de la recherche. Vous n'aurez pas à justifier votre décision.

Si vous décidez de ne pas participer à la recherche, vous n'aurez pas à vous justifier et la qualité des soins dont vous devez bénéficier n'en sera pas modifiée.

### **3. CADRE GENERAL ET OBJECTIFS DE LA RECHERCHE**

Vous (ou vos parents) avez déjà, en tant que patient atteint de diabète de type1(DT1) lors d'un précédent prélèvement sanguin, donné votre consentement pour que l'ADN extrait de votre sang soit conservé au Centre de Biologie et de Pathologie Est du Groupement Hospitalier Est dont dépend l'Hôpital Femme-Mère-Enfant et étudié afin de mettre en évidence ou préciser des gènes qui prédisposent au diabète de type 1.

Dans le cadre du présent projet, nous souhaitons étudier le rôle des facteurs génétiques impliqués dans le contrôle de la dose d'insuline requise (DIR) pour obtenir un bon équilibre glycémique et étudier l'évolution de la DIR à 2 ans puis au cours de votre maladie.

Pour ce nouveau projet de recherche nous avons besoin d'un nouveau consentement écrit de votre part.

### **4. DEROULEMENT DE LA RECHERCHE**

Aucun nouveau prélèvement de sang n'est prévu pour cette nouvelle recherche. Seul votre consentement pour cette nouvelle recherche portant sur vos données génétiques et sur les données médicales relatives à votre diabète vous est demandé.

L'analyse des résultats portera sur environ 1000 patients et est prévue pour se dérouler sur une période de 5 ans.

Par ailleurs, aucune indemnisation n'est prévue dans le cadre de cette recherche.

### **5. DROITS D'INFORMATION CONCERNANT VOTRE SANTE**

Vous avez droit d'avoir communication en cours et à l'issue de la recherche des informations concernant votre santé, conformément à l'article L.1122-1 du Code de la Santé Publique.

Si une ou plusieurs anomalies apparaissent dans vos résultats d'analyse, nous vous en informerons et nous vous conseillerons. Ces résultats seront transmis au médecin de votre choix.



## **7. DROIT D'ETRE INFORME DES RESULTATS GLOBAUX**

Vous avez le droit d'être informé(e) des résultats globaux de la recherche à l'issue de celle-ci, conformément au dernier alinéa de l'article L.1122-1 du Code de la Santé Publique, par le Pr Nicolino soit à l'issue des analyses. Les résultats de cette recherche peuvent être présentés à des congrès ou dans des publications scientifiques. Cependant, vos données personnelles ne seront aucunement identifiables car elles auront été préalablement rendues confidentielles grâce à un codage particulier qui ne mentionne ni votre nom, ni votre prénom.

## **8. CONFIDENTIALITE ET TRAITEMENTS DES DONNEES INFORMATISEES**

Dans le cadre de la recherche médicale à laquelle l'Inserm vous propose de participer, un traitement de vos données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de l'objectif de cette dernière.

A cette fin, les données médicales, vous concernant, et les données relatives à vos origines ethniques collectées par le Pr Nicolino seront transmises de manière sécurisée à l'équipe de recherche de l'INSERM UMR-S958 située à la Faculté de Médecine Villemin 10 av de Verdun 75010 Paris. Ces données seront identifiées par un numéro de code spécifique aux analyses génétiques. Ces données pourront également, dans des conditions assurant leur confidentialité, être transmises aux Autorités de Santé Françaises.

Conformément aux dispositions de la loi relative à l'informatique, aux fichiers et aux libertés, vous disposez d'un droit d'accès et de rectification de ces données. Vous disposez également d'un droit d'opposition à la transmission des données couvertes par le secret professionnel susceptibles d'être utilisées dans le cadre de cette recherche et d'être traitées. Vous pouvez également accéder directement ou par l'intermédiaire d'un médecin de votre choix, à l'ensemble de vos données médicales en application des dispositions de l'article L. 1111-7 du Code de la Santé Publique. Ces droits s'exercent auprès du médecin qui vous suit dans le cadre de la recherche et qui connaît votre identité.

## **9. INFORMATIONS SUR LE DEVENIR DE VOTRE ADN A L'ISSUE DE LA RECHERCHE**

### ***En fin de recherche***

Après utilisation d'une fraction d'ADN pour les besoins de cette recherche, l'ADN restant sera conservé dans la banque d'ADN, sous la responsabilité du Pr Nicolino. La banque originale d'ADN est hébergée au Centre de Biologie et de Pathologie Est du Groupement Hospitalier Est dont dépend l'Hôpital Femme-Mère-Enfant. Une déclaration de collection a été faite par ce centre sous la référence DC-2008-72.

En cas de nouvelle recherche programmée, vos données génétiques ne pourront en aucun cas être étudiées sans un nouveau recueil de consentement écrit de votre part.

D'autre part, vous pouvez librement, et à tout moment, sans conséquence pour votre prise en charge médicale, vous opposer à l'utilisation ultérieure de vos données génétiques à but de recherche, en vous adressant au médecin qui vous suit dans le cadre de cette recherche.

## **10. CADRE LEGISLATIF DE LA RECHERCHE MEDICALE**

Cette recherche est réalisée conformément aux dispositions légales en vigueur régissant les recherches médicales. Ainsi cette recherche est réalisée conformément aux dispositions :

- De la loi relative à la bioéthique (loi du 7 juillet 2011 modifiée par la loi du 6 août 2013)
- De la loi relative l'Informatique, aux fichiers et aux libertés (loi LIL du 6 janvier 1978 modifiée par la loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel du 6 août 2004)
- Du Code de la Santé Publique

Cette recherche a reçu l'avis favorable du Comité de Protection des Personnes Ile de France IV le 24 avril 2014 et l'autorisation de la CNIL le 16 février 2015. L'amendement n°1 a reçu l'avis favorable du Comité de Protection des Personnes Ile de France IV le XXX.

Pour pouvoir participer à cette recherche médicale vous devez être affilié(e) à un régime de Sécurité Sociale, à la Couverture Médicale Universelle (CMU) ou à tout régime équivalent.





## 2. RECUEIL DU CONSENTEMENT DU PARTICIPANT ADULTE

Je soussigné(e) \_\_\_\_\_ atteste avoir bien lu et pris connaissance des informations relatives à ma participation à la recherche intitulée « **Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise (DIR) pour obtenir un bon équilibre glycémique dans une cohorte de patients français atteints de diabète de type I (GENDIR)** » exposées par écrit sur les pages précédentes (p 1 à 3) et transmises oralement par le Pr Nicolino en charge de la recherche, et avoir obtenu toutes les réponses aux questions que je lui ai posées.

- J'ai bien été informé(e) que ma participation à cette recherche consiste à donner mon consentement pour étudier mes données génétiques relatives au diabète corrélées à l'évolution de la DIR.
- J'ai été avisé(e) qu'aucune indemnisation n'est prévue pour cette recherche.
- J'ai bien noté que mon droit d'accès à mes données, prévu par la loi du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés, modifiée par la loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel du 6 août 2004 s'exerce à tout moment auprès du médecin qui me suit dans le cadre de la recherche et qui connaît mon identité. Je pourrai exercer mon droit de rectification et d'opposition auprès de ce même médecin
- J'ai eu suffisamment de temps pour réfléchir à ma participation à cette recherche.
- Je suis parfaitement conscient(e) que je peux retirer à tout moment mon consentement de participation à cette recherche quelque soient mes raisons et sans avoir à m'en justifier, sans supporter aucune responsabilité et sans encourir aucun préjudice. J'en informerai simplement le médecin en charge de la recherche.

### **CONSENTEMENT RELATIF AUX DONNEES PERSONNELLES**

J'accepte que les données me concernant, enregistrées à l'occasion de cette recherche, comportant des données médicales, génétiques et données portant sur mes origines ethniques puissent faire l'objet d'un traitement informatique par le gestionnaire ou pour son compte.

J'accepte également que l'ensemble de mon dossier médical puisse être consulté par les personnes habilitées dans le cadre de cette recherche, dans le respect de la confidentialité de mes données et de mon identité.

### **CONSENTEMENT RELATIF A L'ADN**

J'ai été informé(e) que mon ADN reste sous la responsabilité du Pr. Nicolino après la fin de cette recherche. J'accepte par ma signature que les analyses génétiques soient effectuées sur mon ADN si elles n'ont pas encore été effectuées afin de répondre à l'objectif de cette recherche. Je serais éventuellement sollicité(e) pour signer un nouveau consentement dans le cadre d'une nouvelle recherche portant sur mes données génétiques liées à mon diabète. Ces recherches seront menées par l'équipe de recherche de l'INSERM UMR-S958 située à la Faculté de Médecine Villemin, 10 av de Verdun 75010 Paris.

**J'accepte librement et volontairement de participer à la recherche décrite ci-dessus.** Mon consentement ne décharge en rien le médecin en charge de la recherche et le gestionnaire de l'ensemble de leurs responsabilités et je conserve tous mes droits garantis par la loi.

<b>Personne donnant son consentement</b> : NOM, Prénom : _____ Date : _____ Signature : _____	
<b>Médecin en charge de la recherche</b> : NOM, Prénom : _____ Date : _____ Signature : _____	☎ : _____

**La dernière page de ce document doit comporter les signatures de l'investigateur et de la personne sollicitée et doit être datée de la main de la personne qui a consenti dans l'emplacement prévu et qui lui est réservé. Le paraphe n'engage en rien la personne participante et son recueil ne peut lui être imposé.**

*Ce document d'information et de consentement est établi en deux exemplaires originaux : **un exemplaire est remis à la personne**, un exemplaire est conservé pendant la durée légale de conservation des documents de la recherche, par le médecin en charge de la recherche.*

**Annexe 4 :** Cahier de recueil de données (CRF) de l'étude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une collection de patients atteints de diabète de type I (GENDIR)



**Etude des facteurs génétiques impliqués dans le contrôle de la dose d'insuline journalière requise pour obtenir un bon équilibre glycémique dans une collection de patients atteints de diabète de type I (GENDIR)**

**CAHIER DE RECUEIL DE DONNEES**

Code Recherche	L 00   _   _   _   _   -   _   _   _   <i>n° famille                      n° sujet</i>
Sexe	F   -   M
Date de naissance	_   _   -   _   _   _   _   <i>mois                      année</i>
Date de diagnostic	_   _   -   _   _   -   _   _   _   _
Date de recueil du consentement	_   _   -   _   _   -   _   _   _   _

Responsable scientifique : Flavie Mathieu

INSERM UMR-S958 - Génétique des Diabètes  
 Université Paris Diderot - Paris 7.  
 Faculté de médecine Villemin  
 10, avenue de Verdun. 75010 PARIS  
 Tel : 01 57 27 86 71  
 E-mail : Flavie.Mathieu@inserm.fr

Gestionnaire : INSERM



## SOMMAIRE

MODALITES DE REMPLISSAGE DES CAHIERS DE RECUEIL DE DONNEES .....	3
CRITÈRES DE SÉLECTION .....	4
CARACTÉRISTIQUES SOCIO-DÉMOGRAPHIQUES.....	5
CONSANGUINITE .....	5
NIVEAU D'ETUDES DU PERE.....	5
NIVEAU D'ETUDES DE LA MERE .....	5
NIVEAU D'ETUDES DE L'ENFANT .....	5
PARAMETRES PHYSIQUES .....	7
PRESSION SANGUINE (COUCHÉ).....	8
STATUT PUBERTAIRE (CLASSIFICATION DE TANNER).....	8
HISTOIRE PÉRI ET NÉONATALE .....	9
HISTOIRE MEDICALE PETITE ENFANCE.....	9
SUPPLEMENTATION EN VITAMINE D.....	9
MALADIES INFECTIEUSES .....	10
MALADIES ALLERGIQUES ET INFLAMMATOIRES- VIE ENTIERE .....	10
MALADIES AUTO-IMMUNES- VIE ENTIERE .....	11
RECHERCHE D'AUTO-ANTICORPS .....	11
ETUDE METABOLIQUE.....	13
MODE DE DEBUT .....	14
COMPLICATIONS DEGENERATIVES DU DIABETE.....	14
ANTÉCÉDENTS FAMILIAUX.....	14
ANTECEDENTS FAMILIAUX CHEZ UN APPARENTE DE PREMIER DEGRE (PERE, MERE, FRERE, SŒUR) .....	14
TRAITEMENT PAR INJECTION D'INSULINE.....	16
TRAITEMENTS CONCOMITANTS .....	17
ECHANTILLONS ET GENOTYPAGE .....	17
ATTESTATION DE RECUEIL DE DONNEES .....	17



### MODALITES DE REMPLISSAGE DES CAHIERS DE RECUEIL DE DONNEES

- Utiliser un stylo à bille noir ou bleu
- Ne pas utiliser d'abréviation, mais écrire en toute lettre
- Utiliser un langage médical standard
- Valeurs numériques :
  - cadrer les valeurs numériques à droite
  - ne pas ajouter de virgules, elles sont précodées si nécessaire
  - ne pas laisser de cases vides, mettre un zéro

Incorrect    | 2 | 1 |    Correct    | 0 | 2 | 1 |

- Dates : enregistrer les dates sous la forme | | | | - | | | | - | | | | | | | |  
J J - M M - A A A A

- En cas d'absence de données utiliser et inscrire les codes suivants :

- NA : non applicable
- NF : non fait
- NC : non connu
- DM : donnée manquante

Exemple : poids    | | D | M |    Kg

- Si une donnée doit être corrigée, rayer l'information erronée  
*(elle doit, toutefois, rester lisible)*

- remplacer par l'information correcte
- **dater et parapher la correction**

Exemple : poids    <sup>0 7 8</sup>  
~~10 7 0~~    Kg

**Ne jamais utiliser de blanc correcteur**

Etude <b>GENDIR</b>	Code du sujet XXX       -       -       n° centre n° famille n° sujet	DONNEES CLINIQUES	page 4 / 17
------------------------	---	-------------------	----------------

**CRITERES DE SELECTION****CRITERES D'INCLUSION -**

Critères	OUI	NON
Etre patient diabétique de type 1	<input type="checkbox"/>	<input type="checkbox"/>
Avoir un échantillon présent dans la collection hébergée au Centre de Biologie et de Pathologie Est du Groupement Hospitalier Est dont dépend l'Hôpital Femme-Mère-Enfant déclarée sous la référence DC-2011-1437.	<input type="checkbox"/>	<input type="checkbox"/>
Avoir un âge de moins de 18 ans	<input type="checkbox"/>	<input type="checkbox"/>
Avoir une valeur de la DIR et un dosage d'HbA1c après 2 ans de maladie	<input type="checkbox"/>	<input type="checkbox"/>

***TOUS*** les critères d'**INCLUSION** doivent être cotés **OUI**,

**CRITERES DE NON-INCLUSION**

Critères	OUI	NON
Avoir un apparenté inclus dans l'étude	<input type="checkbox"/>	<input type="checkbox"/>

***Le*** critère de **NON-INCLUSION** doit être coté **NON**



Etude <b>GENDIR</b>	Code du sujet XXX       -       -       n° centre      n° famille      n° sujet	DONNEES CLINIQUES	page 5 / 17
------------------------	---	-------------------	----------------

**CARACTERISTIQUES SOCIO-DEMOGRAPHIQUES****LIEU DE NAISSANCE :**

Pays : Si né en France, département : ..... | | |

Si né hors de France, pays : .....

Date d'arrivée en France : | | | / | | | / | | | | | |

**ORIGINE GEOGRAPHIQUE DES PARENTS BIOLOGIQUES :**

	Département	PAYS
Père :	.....	.....
Mère :	.....	.....

**CONSANGUINITE**Consanguinité dans la famille :  OUI NON NC

Si oui, spécifier le degré : (cousins germains, cousins issus de cousins germains, ...) \_\_\_\_\_

**NIVEAU D'ETUDES DU PERE**

Scolarité primaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Scolarité secondaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
CAP/BEP	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Brevet des collèges	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Baccalauréat	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Etudes supérieures	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Niveau maximal atteint : (Mettre la classe entre CP à DOCTORAT)			

**NIVEAU D'ETUDES DE LA MERE**

Scolarité primaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Scolarité secondaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
CAP/BEP	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Brevet des collèges	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Baccalauréat	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Etudes supérieures	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Niveau maximal atteint : (Mettre la classe entre CP à DOCTORAT)			

**NIVEAU D'ETUDES DE L'ENFANT AU DIAGNOSTIC**

Scolarité primaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Scolarité secondaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
CAP/BEP	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Brevet des collèges	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Baccalauréat	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Etudes supérieures	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Niveau maximal atteint : (Mettre la classe entre CP à DOCTORAT)			

<b>Etude</b> <b>GENDIR</b>	<b>Code du sujet</b> XXX       -         -         <small>n° centre      n° famille      n° sujet</small>	<i>DONNEES CLINIQUES</i>	<b>page</b> <b>6 / 17</b>
-------------------------------	---	--------------------------	------------------------------

**NIVEAU D'ETUDES DE L'ENFANT A 2 ANS APRES LE DIAGNOSTIC**

Scolarité primaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Scolarité secondaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
CAP/BEP	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Brevet des collèges	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Baccalauréat	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Etudes supérieures	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Niveau maximal atteint : <i>(Mettre la classe entre CP à DOCTORAT)</i>			

**NIVEAU D'ETUDES DE L'ENFANT A 5 ANS APRES LE DIAGNOSTIC**

Scolarité primaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Scolarité secondaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
CAP/BEP	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Brevet des collèges	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Baccalauréat	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Etudes supérieures	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Niveau maximal atteint : <i>(Mettre la classe entre CP à DOCTORAT)</i>			

**NIVEAU D'ETUDES DE L'ENFANT A 10 ANS APRES LE DIAGNOSTIC**

Scolarité primaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Scolarité secondaire achevée	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
CAP/BEP	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Brevet des collèges	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Baccalauréat	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Etudes supérieures	<input type="checkbox"/> Oui	<input type="checkbox"/> Non	<input type="checkbox"/> NC
Niveau maximal atteint : <i>(Mettre la classe entre CP à DOCTORAT)</i>			





<b>Etude</b> <b>GENDIR</b>	<b>Code du sujet</b> XXX       -       -       <small>n° centre      n° famille      n° sujet</small>	<b>DONNEES CLINIQUES</b>	<b>page</b> <b>9 / 17</b>
-------------------------------	---	--------------------------	------------------------------

**HISTOIRE PERI ET NEONATALE**

Age de la mère à la naissance du sujet : | | | ans

Age du père à la naissance du sujet : | | | ans

Naissance :

 Par voie Basse       Par césarienne       NC

Rang de naissance : | | |

Taille à la naissance : | | | | cm

Poids à la naissance : | | kg | | | | g

Périmètre crânien à la naissance : | | | cm

Terme gestationnel : | | | Semaines

Hospitalisation en néonatalogie :	<input type="checkbox"/> OUI	<input type="checkbox"/> NON	<input type="checkbox"/> NC
Terme:	<input type="checkbox"/> Pré maturité	<input type="checkbox"/> Né à terme	<input type="checkbox"/> Post-maturité
Hypoglycémie néonatale :	<input type="checkbox"/> OUI	<input type="checkbox"/> NON	<input type="checkbox"/> NC
Ictère :	<input type="checkbox"/> OUI	<input type="checkbox"/> NON	<input type="checkbox"/> NC
Infection materno-fœtale :	<input type="checkbox"/> OUI	<input type="checkbox"/> NON	<input type="checkbox"/> NC
Stress néonatal :	<input type="checkbox"/> OUI	<input type="checkbox"/> NON	<input type="checkbox"/> NC

Groupe sanguin de la mère du sujet :  A  B  AB  O Rhésus :  Négatif  PositifGroupe sanguin du sujet :  A  B  AB  O Rhésus :  Négatif  Positif**HISTOIRE MEDICALE PETITE ENFANCE****SUPPLEMENTATION EN VITAMINE D**
 OUI       NON       NE SAIS PAS
Si oui, préciser : Date de prise | | | | | | | | | | | | | |  
J J / M M / A A A A

Si oui, préciser : Date de prise | | | | | | | | | | | | | |

Si oui, préciser : Date de prise | | | | | | | | | | | | | |

Si oui, préciser : Date de prise | | | | | | | | | | | | | |

Si oui, préciser : Date de prise | | | | | | | | | | | | | |

Si oui, préciser : Date de prise | | | | | | | | | | | | | |

Si oui, préciser : Date de prise | | | | | | | | | | | | | |

Etude <b>GENDIR</b>	Code du sujet XXX         -         -         n° centre n° famille n° sujet	DONNEES CLINIQUES	page 10 / 17
------------------------	---	-------------------	-----------------

**MALADIES INFECTIEUSES**

Maladie infectieuse :  OUI  NON  NE SAIS PAS

Si oui, préciser : Infection : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Si oui, préciser : Infection : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Si oui, préciser : Infection : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

**Hépatite virale**  oui  non  NE SAIS PAS

Si oui, préciser : Type  A  B  C  D  E  non classée

Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Traitée  oui  non

Equilibrée  oui  non

**Mononucléose infectieuse (virus d'Epstein-Barr) :**  oui  non  NE SAIS PAS

Si oui, Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

**MALADIES ALLERGIQUES ET INFLAMMATOIRES- VIE ENTIERE**

Pathologies allergiques et inflammatoires  oui  non Si oui spécifier :

Si oui, préciser : Pathologie : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Si oui, préciser : Pathologie : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Si oui, préciser : Pathologie : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Si oui, préciser : Pathologie : \_\_\_\_\_ Date de diagnostic |\_|\_|\_|\_|\_|\_|\_|\_|\_|

Etude <b>GENDIR</b>	Code du sujet XXX         -         -         n° centre n° famille n° sujet	DONNEES CLINIQUES	page 11 / 17
------------------------	---	-------------------	-----------------

**MALADIES AUTO-IMMUNES- VIE ENTIERE**

Maladie auto-immune:  oui  non  NE SAIS PAS

Si oui, préciser :

Maladie : \_\_\_\_\_ Date de diagnostic |\_|\_|/|\_|\_|/|\_|\_| Traitement  oui  non

Maladie : \_\_\_\_\_ Date de diagnostic |\_|\_|/|\_|\_|/|\_|\_| Traitement  oui  non

Maladie : \_\_\_\_\_ Date de diagnostic |\_|\_|/|\_|\_|/|\_|\_| Traitement  oui  non

Maladie : \_\_\_\_\_ Date de diagnostic |\_|\_|/|\_|\_|/|\_|\_| Traitement  oui  non

**RECHERCHE D'AUTO-ANTICORPS**

Anticorps spécifiques	Réalisé	Date(s) de dosage(s)	Valeur	Unité	Normes du laboratoire
Glutamic Acid Decarboxylase (GAD65)	<input type="checkbox"/> OUI <input type="checkbox"/> NON	_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
Thyrosine phosphatase insulina antigen (IA2)	<input type="checkbox"/> OUI <input type="checkbox"/> NON	_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
Zinc transporter 8 (ZnT8A)	<input type="checkbox"/> OUI <input type="checkbox"/> NON	_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
INS	<input type="checkbox"/> OUI <input type="checkbox"/> NON	_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
ICA	<input type="checkbox"/> OUI <input type="checkbox"/> NON	_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
Autre (Spécifier) _____ _____ _____	<input type="checkbox"/> OUI <input type="checkbox"/> NON	_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml
		_ _ / _ _ / _ _		UI/ml	UI/ml





<b>Etude</b> <b>GENDIR</b>	<b>Code du sujet</b> XXX       -       -       n° centre      n° famille      n° sujet	<i>DONNEES CLINIQUES</i>	<b>page</b> 14 / 17
-------------------------------	--	--------------------------	------------------------

**MODE DE DEBUT**

- Réanimation  OUI  NON  NC
- Acidocétose  OUI  NON  NC
- Cétose (pH≤7.20+réserve alcaline ≤12)  OUI  NON  NC
- Polyuro-Polydipsie  OUI  NON  NC
- Perte de Poids  OUI  NON  NC si oui de |\_|\_| kg en |\_|\_| mois
- Autre symptôme  OUI  NON  NC Préciser \_\_\_\_\_

**COMPLICATIONS DEGENERATIVES DU DIABETE**

- Complications  OUI  NON  NC
- Si oui :
- Rétinopathie  OUI  NON  NC Age de début : |\_|\_| ans
- Néphropathie  OUI  NON  NC Age de début : |\_|\_| ans
- Neuropathie  OUI  NON  NC Age de début : |\_|\_| ans
- Maladie Cardiovasculaire  OUI  NON  NC Age de début : |\_|\_| ans
- Greffe  OUI  NON Type : \_\_\_\_\_ Age : |\_|\_| ans
- Autre \_\_\_\_\_  OUI  NON  NC Age de début : |\_|\_| ans

**ANTECEDENTS FAMILIAUX****ANTECEDENTS FAMILIAUX CHEZ UN APPARENTE DE PREMIER DEGRE (PERE, MERE, FRERE, SŒUR)**• **ANTECEDENT DE DIABETE OU OBESITE CHEZ LE PERE**

Année de naissance du père : |\_|\_|\_|\_|

- Diabète de Type 1
- Diabète de Type 2
- Diabète MODY
- Pré-diabète de type 1
- Diabète non spécifié
- Obésité  OUI  NON  NC
- Hyperlipidémie  OUI  NON  NC
- Pathologie Auto-immune  OUI  NON  NC Si oui Spécifier \_\_\_\_\_
- Autre pathologie (Spécifier) : \_\_\_\_\_





Etude <b>GENDIR</b>	Code du sujet XXX       -           -           <small>n° centre      n° famille      n° sujet</small>	DONNEES CLINIQUES	page 17 / 17
------------------------	--	-------------------	-----------------

**TRAITEMENTS CONCOMITANTS**

Traitement (nom commercial)	Indication thérapeutique	Date de début JJ / MM / AAAA	Date de fin JJ / MM / AAAA
		_ _  /  _ _  /  _ _ _ _	_ _  /  _ _  /  _ _ _ _
		_ _  /  _ _  /  _ _ _ _	_ _  /  _ _  /  _ _ _ _
		_ _  /  _ _  /  _ _ _ _	_ _  /  _ _  /  _ _ _ _
		_ _  /  _ _  /  _ _ _ _	_ _  /  _ _  /  _ _ _ _
		_ _  /  _ _  /  _ _ _ _	_ _  /  _ _  /  _ _ _ _

**ECHANTILLONS ET GENOTYPAGE**

TYPAGE HLA :                      HLA-DR    |\_|\_| / |\_|\_|                      HLA-DQB1    |\_|\_| / |\_|\_|

ADN DISPONIBLE :                       OUI     NON

LIGNEES DISPONIBLES:                       OUI     NON

GENOTYPAGE :                       EFFECTUE     A REALISER

**ATTESTATION DE RECUEIL DE DONNEES**

Observations éventuelles:  
 .....  
 .....  
 .....

Recueil des données effectué le :

**Nom du médecin de l'étude**

Date : |\_|\_| - |\_|\_| - |\_|\_|\_|\_|\_|  
    J J    M M    A A A A

Signature :

## Annexe 5 : Phénotype et processus biologiques associés aux gènes testés dans l'étude GENDIR

Gene	Associations in the EBI GWAS catalog*	Biological Process (GO) **	Classification
BNC2	adolescent idiopathic scoliosis (26211971; 30395268; 31417091); arterial stiffness measurement (31235810); basal cell carcinoma (27539887; 31174203); <b>BMI-adjusted waist circumference (28552196); body fat distribution (30664634); body height (30595370); body mass index (30595370; 30239722)</b> ; carrier status, ovarian carcinoma (28346442); cutaneous squamous cell carcinoma (27424798); Eczema (30595370); Eczema, allergic rhinitis (31361310); <b>eosinophil count (30595370)</b> ; facial pigmentation measurement (25705849); freckles (20585627; 29895819); FEV/FEC ratio (30595370); hair color (30531825); hair colour measurement (29662168; 30595370); heel bone mineral density (30048462; 328869591; 0598549; 30595370); Heschl's gyrus morphology measurement (25130324); <b>high density lipoprotein cholesterol measurement (29507422); leukocyte count (30595370)</b> ; lobe attachment (29198719); <b>low density lipoprotein cholesterol measurement (29507422); low density lipoprotein cholesterol measurement, coronary artery disease (30525989)</b> ; Malignant epithelial tumor of ovary (25581431); non-melanoma skin carcinoma (29739929); ovarian carcinoma (19648919; 23544013; 23535730; 28346442; 28346442); pulse pressure measurement (30224653); pyroglutamine measurement (23934736); reaction time measurement (29844566); response to bronchodilator, chronic obstructive pulmonary disease, FEV/FEC ratio (26634245); response to diisocyanate, asthma (25918132); <b>reticulocyte count (27863252)</b> ; risk-taking behaviour (30643258); squamous cell carcinoma (26829030; 31174203); skin pigmentation (30166351); smoking status measurement (30643251); sunburn (30595370); suntan (29739929); <b>total cholesterol measurement (29083408; 29507422; 30926973)</b> ; velopharyngeal dysfunction (29855589); <b>visceral adipose tissue measurement (31501611)</b> ; white matter microstructure measurement (31666681);	/	<b>Metabolism; Immunity</b>
C18orf16	Heart rate in heart failure with reduced ejection fraction (31113495)	/	Other

CCRS	Blood protein levels in cardiovascular risk (28369058) ; <b>Blood protein levels (28915241)</b> ; <b>Granulocyte percentage of myeloid white cells (27863252)</b>	<p>antimicrobial humoral immune response mediated by antimicrobial peptide; cell-cell signaling; cellular calcium ion homeostasis; <b>cellular response to interferon-gamma</b>; <b>cellular response to interleukin-1</b>; <b>cellular response to lipopolysaccharide</b>; <b>cellular response to tumor necrosis factor</b>; chemokine-mediated signaling pathway; cytokine-mediated signaling pathway; <b>eosinophil chemotaxis</b>; G protein-coupled receptor signaling pathway; <b>immune response</b>; <b>inflammatory response</b>; interspecies interaction between organisms; <b>killing of cells of other organism</b>; <b>leukocyte chemotaxis</b>; <b>lymphocyte chemotaxis</b>; modification of morphology or physiology of other organism; <b>monocyte chemotaxis</b>; <b>negative regulation by host of viral transcription</b>; <b>negative regulation of macrophage apoptotic process</b>; negative regulation of multi-organism process; <b>neutrophil chemotaxis</b>; positive regulation of calcium ion transport; positive regulation of cell migration; <b>positive regulation of cellular metabolic process</b>; positive regulation of cellular process; positive regulation of cytosolic calcium ion concentration; positive regulation of ERK1 and ERK2 cascade; positive regulation of GTPase activity; <b>positive regulation of immune system process</b>; positive regulation of intracellular signal transduction; <b>positive regulation of leukocyte chemotaxis</b>; <b>positive regulation of lymphocyte migration</b>; <b>positive regulation of macromolecule metabolic process</b>; positive regulation of molecular function; <b>positive regulation of natural killer cell chemotaxis</b>; <b>positive regulation of nitrogen compound metabolic process</b>; positive regulation of transport; regulation of anatomical structure morphogenesis; regulation of catalytic activity; <b>regulation of cellular protein metabolic process</b>; regulation of localization; <b>regulation of lymphocyte chemotaxis</b>; regulation of protein phosphorylation; regulation of response to external stimulus; regulation of sequestering of calcium ion; regulation of signaling receptor activity; regulation of symbiosis, encompassing mutualism through parasitism; response to lipopolysaccharide; response to other organism; response to virus ; T cell chemotaxis</p>	<b>Metabolism;</b> <b>Immunity</b>
------	---	--	---------------------------------------

CTLA4	<p><b>Alopecia areata (20596022; 25608926); Celiac disease (20190752); Graves' disease (21841780); Hypothyroidism (27182965); Inflammatory bowel disease (26192919); Myasthenia gravis (25643325); Primary biliary cholangitis (28425483); Rheumatoid arthritis (19503088; 20453842; 24390342); Subcutaneous adipose tissue (29178545); Tonsillectomy (28928442); Type 1 diabetes (18978792; 19430480; 21829393); Ulcerative colitis (26192919); Vitiligo (27723757)</b></p>	<p><b>T cell costimulation; Fc-epsilon receptor signaling pathway; Immune response-regulating cell surface receptor signaling pathway; Positive regulation of lymphocyte activation; Immune response-regulating signaling pathway; Antigen receptor-mediated signaling pathway; Positive regulation of leukocyte proliferation; phosphatidylinositol-mediated signaling; fibroblast growth factor receptor signaling pathway; positive regulation of T cell activation; regulation of lymphocyte proliferation; cellular response to fibroblast; growth factor stimulus; epidermal growth factor receptor signaling pathway; positive regulation of lymphocyte proliferation; B cell receptor signaling pathway; neurotrophin TRK receptor signaling pathway; immune response-activating signal transduction; immune response; positive regulation of cell differentiation; regulation of immune response; cellular response to oxygen-containing compound</b></p>	<p><b>Autoimmune-disease ; Immunity</b></p>
DPP6	<p>age at onset, Myopia (29808027); alcohol consumption measurement (30643258; 30643251; 31358974; 30679032); amyotrophic lateral sclerosis (18057069; 18987618; 18084291) ; attention deficit hyperactivity disorder (23527680); body height (28552196); bone mineral accretion measurement, response to corticosteroid, asthma (26025128); cognitive decline measurement (30954325); colorectal cancer, diet measurement (24743840); DNA methylation (30348214); forced expiratory volume, response to bronchodilator, chronic obstructive pulmonary disease (26634245); gestational age, birth measurement (27490719) ; lean body mass (28552196); ligament rupture (28358823); movement disorder (21826085); nasopharyngeal neoplasm (24376627); pancreatic carcinoma (20686608); pelvic organ prolapse (26545240); phosphatidylcholine 38:3 measurement (23823483); post-traumatic stress disorder (24677629); QT interval (23534349) ; risk-taking behavior (30643258); serum IgG glycosylation measurement (23382691); smoking cessation (30643251); smoking status measurement (30595370; 30643251); social interaction measurement (29970889) ; thalamus volume (31530798); velopharyngeal dysfunction (29855589); venous thromboembolism (28203683)</p>	<p>action potential ; cardiac muscle cell action potential ; inorganic cation transmembrane transport ; metal ion transport ; monovalent inorganic cation transport ; potassium ion transmembrane transport ; protein homooligomerization; regulation of cation transmembrane transport; regulation of ion transmembrane transport; regulation of metal ion transport; regulation of potassium ion transmembrane transport</p>	<p>Other</p>
ERBB3	<p><b>Alopecia areata (25608926); Anorexia nervosa (28494655); Cognitive function (25644384); Educational attainment (25201988); Hypothyroidism (27182965); Polycystic ovary syndrome (26416764); Psoriasis (20953189); Type 1 diabetes (17554260; 17554300; 18198356; 18978792; 19430480; 21829393); Vitiligo (22951725); Obesity-related traits (23251661)</b></p>	<p>fibroblast growth factor receptor signaling pathway; Fc-epsilon receptor signaling pathway; cellular response to fibroblast growth factor stimulus; epidermal growth factor receptor signaling pathway; neurotrophin TRK receptor signaling pathway; phosphatidylinositol-mediated signaling; <b>innate immune response;</b> cellular response to growth factor stimulus; wound healing; transmembrane receptor protein tyrosine kinase signaling pathway; positive regulation of protein kinase activity; activation of protein kinase activity; regulation of protein tyrosine kinase activity; regulation of protein phosphorylation; <b>defense response;</b> blood coagulation; <b>cellular response to endogenous stimulus;</b> positive regulation of protein tyrosine kinase activity; intracellular signal transduction; <b>insulin receptor</b></p>	<p><b>Autoimmune-disease; Metabolism; Immunity</b></p>

		<b>signaling pathway; immune system process</b> ; activation of transmembrane receptor protein tyrosine kinase activity; regulation of intracellular signal transduction; MAPK cascade; <b>cellular response to insulin stimulus</b>	
FANCI	Age at menopause (30595370); Height (25282103); Red cell distribution width (30595370)	DNA repair; interstrand cross-link repair; meiotic nuclear division	Other
FLJ23577	/	cell projection organization; cilium assembly; cilium movement; epithelial cilium movement; motile cilium assembly	Other
IL6	C-reactive protein (21196492); Cerebrospinal fluid clusterin levels in APOEε4- carriers (26545630); Dialysis-related mortality (21546767); Fibrinogen (23969696); Longevity (26912274)	<b>response to cytokine</b> ; response to lipopolysaccharide; <b>cellular response to cytokine stimulus</b> ; <b>cellular response to interleukin-6</b> ; cellular response to organic substance; <b>interleukin-6-mediated signaling pathway</b> ; response to organonitrogen compound; response to lipid; <b>response to bacterium</b> ; response to hormone; cellular response to hormone stimulus; response to glucocorticoid; <b>cellular response to endogenous stimulus</b> ; response to purine-containing compound; positive regulation of defense response; cytokine-mediated signaling pathway; response to steroid hormone; response to organic cyclic compound; positive regulation of cell differentiation; <b>regulation of immune response</b> ; defense response; response to nitrogen compound; multi-organism process; aging; regulation of cell differentiation; <b>regulation of inflammatory response</b> ; <b>response to organic substance</b> ; <b>positive regulation of leukocyte differentiation</b> ; response to organophosphorus; response to alcohol; acute-phase response; <b>positive regulation of macromolecule metabolic process</b> ; <b>positive regulation of cytokine production</b> ; <b>positive regulation of cellular metabolic process</b> ; <b>regulation of innate immune response</b> ; positive regulation of smooth muscle cell proliferation; <b>positive regulation of response to stimulus</b> ; positive regulation of cellular process; regulation of defense response; <b>immune response</b> ; cell surface receptor signaling pathway; <b>inflammatory response</b> ; positive regulation of cell proliferation; response to peptide hormone; positive regulation of multicellular organismal process; <b>regulation of immune system process</b> ; regulation of apoptotic process; positive regulation of transcription, DNA-templated; <b>regulation of response to external stimulus</b> ; regulation of cell death	<b>Immunity</b>



LPPR4	/	CDP-diacylglycerol biosynthetic process; CDP-diacylglycerol metabolic process; cellular biosynthetic process; <b>cellular lipid metabolic process</b> ; glycerolipid biosynthetic process; <b>glycerolipid metabolic process</b> ; <b>glycerophospholipid metabolic process</b> ; lipid biosynthetic process; <b>nucleotide metabolic process</b> ; <b>organophosphate metabolic process</b> ; <b>phosphate-containing compound metabolic process</b> ; phosphatidic acid biosynthetic process; <b>phosphatidic acid metabolic process</b> ; phosphatidylethanolamine acyl-chain remodeling; phospholipid biosynthetic process; <b>phospholipid metabolic process</b>	Metabolism
-------	---	---	------------

LRP5	<p>appendicular lean mass (31761296); bone density (18455228; 19801982; 22504420; 24249740; 29304378); bone fracture (30158200); Drugs affecting bone structure and mineralization use measurement (31015401); femoral neck bone mineral density (29499414); hair colour measurement (30595370); heel bone mineral density (30598549; 30595370; 28869591; 30048462); mathematical ability (30038396); <b>mean corpuscular hemoglobin (30595370)</b>; spine bone mineral density (29499414); synophrys measurement (27182965)</p>	<p>anatomical structure formation involved in morphogenesis; anatomical structure morphogenesis; animal organ development; animal organ morphogenesis; anterior/posterior pattern specification; axis specification; beta-catenin destruction complex disassembly; brain development; branching morphogenesis of an epithelial tube; canonical Wnt signaling pathway; canonical Wnt signaling pathway involved in heart development; canonical Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation; canonical Wnt signaling pathway involved in negative regulation of apoptotic process; cell development; cell differentiation; cell fate commitment; cell morphogenesis involved in differentiation; cellular component disassembly; cellular response to retinoic acid; central nervous system neuron differentiation; chordate embryonic development; circulatory system development; developmental growth involved in morphogenesis; embryonic axis specification; embryonic digit morphogenesis; embryonic epithelial tube formation; embryonic limb morphogenesis; embryonic morphogenesis; embryonic organ morphogenesis; epithelial tube morphogenesis; epithelium development; extracellular matrix-cell signaling; eye development; eye morphogenesis; forebrain anterior/posterior pattern specification; forebrain development; forebrain generation of neurons; formation of primary germ layer; gastrulation; gastrulation with mouth forming second; generation of neurons; gland development; inner ear morphogenesis; mesoderm formation; multicellular organismal process; muscle cell differentiation; muscle structure development; negative regulation of BMP signaling pathway; negative regulation of canonical Wnt signaling pathway; negative regulation of cell death; negative regulation of cell differentiation; <b>negative regulation of cellular metabolic process</b>; negative regulation of cellular process; negative regulation of developmental process; negative regulation of embryonic development; negative regulation of fat cell differentiation; negative regulation of neuron death; negative regulation of nitrogen compound metabolic process; negative regulation of oxidative stress-induced neuron death; negative regulation of response to stimulus; negative regulation of signal transduction; neural tube development; neuron differentiation; non-canonical Wnt signaling pathway; ossification; planar cell polarity pathway involved in neural tube closure; positive regulation of canonical Wnt signaling pathway; positive regulation of cell differentiation; positive regulation of cell projection organization; positive regulation of cellular component organization; <b>positive regulation of cellular metabolic process</b>; positive regulation of dermatome development; positive regulation of developmental process; positive regulation of DNA-binding transcription factor activity; positive regulation of embryonic development; positive regulation of gene expression; positive regulation of JUN kinase activity; positive regulation of</p>	Metabolism
------	--	---	------------

		<p>MAPK cascade; positive regulation of mesenchymal cell proliferation; positive regulation of molecular function; positive regulation of multicellular organismal process; positive regulation of neurogenesis; positive regulation of neuron differentiation; positive regulation of neuron projection arborization; positive regulation of neuron projection development; positive regulation of nitrogen compound metabolic process; positive regulation of protein kinase activity; positive regulation of protein phosphorylation; positive regulation of RNA metabolic process; positive regulation of signal transduction; positive regulation of transcription, DNA-templated; positive regulation of transferase activity; positive regulation of Wnt signaling pathway; presynapse assembly; regulation of animal organ formation; regulation of animal organ morphogenesis; regulation of binding; regulation of canonical Wnt signaling pathway; regulation of cell death; regulation of cell differentiation; regulation of cellular component organization; regulation of cellular response to growth factor stimulus; regulation of cellular response to stress; regulation of developmental process; regulation of embryonic development; regulation of establishment of planar polarity; regulation of fat cell differentiation; regulation of heart morphogenesis; regulation of molecular function; regulation of morphogenesis of an epithelium; regulation of multicellular organismal development; regulation of multicellular organismal process; regulation of neuron death; regulation of ossification; regulation of peptidyl-serine phosphorylation; regulation of plasma membrane bounded cell projection organization; regulation of protein binding; regulation of protein kinase activity; regulation of protein phosphorylation; regulation of protein serine/threonine kinase activity; regulation of RNA metabolic process; regulation of signal transduction; regulation of skeletal muscle tissue development; regulation of timing of anagen; regulation of transcription by RNA polymerase II; regulation of transcription, DNA-templated; regulation of transferase activity; response to chemical; response to oxygen-containing compound; response to retinoic acid; retinal blood vessel morphogenesis; secondary palate development; sensory organ development; sensory organ morphogenesis; skeletal system development; somitogenesis; stem cell proliferation; synapse organization; tissue morphogenesis; tube morphogenesis; Wnt signaling pathway; Wnt signaling pathway involved in heart development; Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation; Wnt signaling pathway, calcium modulating pathway; Wnt signaling pathway, planar cell polarity pathway</p>	
MEAT6	/	/	Unknown

PHLDB3	Breast cancer (29059683)	/	Other
PTPN22	<b>Alopecia areata (25608926); Antineutrophil cytoplasmic antibody-associated vasculitis (28029757); Breast cancer (23535729); Crohn's disease (18587394; 21102463; 23128233; 28067908); Gut microbiota (27694959); Hypothyroidism (22493691; 27182965); Immun cells numbers (27863252); Myasthenia gravis (23055271; 26562150); Obesity-related traits (23251661); Pediatric autoimmune diseases (26301688); Rheumatoid arthritis (17554300; 17804836; 18794853; 19503088; 20453842; 21156761; 24390342; 24449572); Systemic lupus erythematosus (19838195; 26502338; 27193031; 27399966); Thyroid peroxidase antibody levels (24586183); Type 1 diabetes and autoimmune thyroid diseases (25936594); Type 1 diabetes (17554260; 17554300; 19430480; 17632545; 18978792; 21829393); Vitiligo (20410501; 27723757)</b>	<b>T cell receptor signaling pathway; T cell costimulation; Regulation of immune response</b>	<b>Autoimmune-disease ; Immunity</b>
RAF1	body height (30595370) ; cardiac hypertrophy (21348951) ; <b>erythrocyte count (30595370) ; hair colour measurement (30595370) ; mosquito bite reaction itch intensity measurement (28199695) ; total cholesterol measurement (20686565; 24097068)</b>	intracellular signal transduction; MAPK cascade; positive regulation of catalytic activity; positive regulation of MAP kinase activity; positive regulation of MAPK cascade; positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway; positive regulation of protein kinase activity; positive regulation of protein localization to membrane; positive regulation of response to stimulus; positive regulation of signal transduction; Ras protein signal transduction; regulation of apoptotic process; regulation of catalytic activity; regulation of cellular localization; regulation of cellular protein metabolic process; regulation of intracellular transport; regulation of phosphate metabolic process; regulation of protein modification process; regulation of response to stress; signal transduction ; stimulatory C-type lectin receptor signaling pathway	<b>Metabolism</b>
RXFP2	Agents acting on the renin-angiotensin system use measurement (31015401); Calcium channel blocker use measurement (31015401); cardiovascular disease (30595370); Diuretic use measurement (31015401); HVA measurement, MHPG measurement (23319000); pulse pressure measurement (30578418); spontaneous preterm birth (31194736); systolic blood pressure (30224653; 30578418); temporomandibular joint disorder (28081371); type I diabetes mellitus, A1C measurement (30674623)	activation of adenylate cyclase activity; adenylate cyclase-activating G protein-coupled receptor signaling pathway; adenylate cyclase-modulating G protein-coupled receptor signaling pathway; G protein-coupled receptor signaling pathway; positive regulation of adenylate cyclase activity; signal transduction	Other
SORCS1	White matter microstructure in first episode schizophrenia (left posterior cingulate cortex) (28924203); Abnormal white matter microstructure in first episode schizophrenia (multivariate analysis) (28924203); Systemic juvenile idiopathic arthritis (27927641); Methadone dose in opioid dependence (28115739); Endometriosis (23472165); Metabolite levels (Pyroglutamine) (23934736); Asparaginase hypersensitivity in acute lymphoblastic leukemia (25987655)	/	Other

SYK	atrial fibrillation (28416822); frontal pole volume measurement (31530798); multiple sclerosis (21833088; 31604244); multiple sclerosis, high density lipoprotein cholesterol measurement (26920376); platelet component distribution width (27863252); platelet count (27863252); prostate carcinoma (22219177); <b>type II diabetes mellitus (22456796)</b> ; vascular dementia (23480133).	<b>antigen receptor-mediated signaling pathway</b> ; cell surface receptor signaling pathway; endocytosis; Fc receptor mediated stimulatory signaling pathway; Fc-epsilon receptor signaling pathway; Fc-gamma receptor signaling pathway; Fc-gamma receptor signaling pathway involved in phagocytosis; <b>immune effector process</b> ; <b>immune response</b> ; <b>immune response-activating cell surface receptor signaling pathway</b> ; <b>immune response-activating signal transduction</b> ; <b>immune system process</b> ; <b>innate immune response-activating signal transduction</b> ; integrin-mediated signaling pathway; intracellular signal transduction; <b>leukocyte activation</b> ; <b>leukocyte activation involved in immune response</b> ; <b>leukocyte differentiation</b> ; <b>lymphocyte activation</b> ; <b>lymphocyte differentiation</b> ; mast cell activation; <b>myeloid leukocyte activation</b> ; phagocytosis; platelet activation; positive regulation of endocytosis; <b>positive regulation of innate immune response</b> ; positive regulation of receptor-mediated endocytosis; positive regulation of response to stimulus; <b>regulation of immune response</b> ; regulation of response to stimulus; response to stress; serotonin secretion by platelet; signal transduction; <b>T cell activation</b> ; <b>T cell receptor signaling pathway</b> ; transmembrane receptor protein tyrosine kinase signaling pathway; transport ; vesicle-mediated transport	Immunity
TLR3	Breast size (22747683)	<b>toll-like receptor signaling pathway</b> ; <b>toll-like receptor 3 signaling pathway</b> ; <b>toll-like receptor 4 signaling pathway</b> ; TRIF-dependent toll-like receptor signaling pathway; <b>positive regulation of innate immune response</b> ; <b>innate immune response</b> ; <b>positive regulation of I-kappaB kinase</b> ; <b>NF-kappaB signaling</b> ; <b>I-kappaB kinase</b> ; <b>NF-kappaB signaling</b> ; <b>toll-like receptor 9 signaling pathway</b> ; <b>positive regulation of type I interferon production</b> ; cytoplasmic pattern recognition receptor signaling pathway; <b>positive regulation of NF-kappaB transcription factor activity</b> ; <b>toll-like receptor TLR1:TLR2 signaling pathway</b> ; <b>toll-like receptor TLR6:TLR2 signaling pathway</b> ; <b>toll-like receptor 2 signaling pathway</b> ; activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway; <b>MyD88-dependent toll-like receptor signaling pathway</b> ; <b>regulation of immune response</b> ; <b>positive regulation of interferon-beta production</b> ; <b>response to other organism</b> ; <b>defense response to other organism</b> ; <b>positive regulation of cytokine production</b> ; <b>defense response to virus</b> ; <b>immune response</b> ; <b>positive regulation of interleukin-6 production</b> ; <b>positive regulation of immune system process</b> ; <b>toll-like receptor 5 signaling pathway</b> ; <b>toll-like receptor 10 signaling pathway</b> ; <b>positive regulation of intracellular signal transduction</b> ; <b>response to bacterium</b>	Immunity
TMEM74	<b>Alopecia (28196072)</b> ; lip morphology measurement (28441456); Reduced ejection fraction, heart failure, heart rate (31113495); t-tau measurement (20932310).	autophagy	Other

TNF	<p><b>AIDS progression (19115949)</b>; Autism spectrum disorder or schizophrenia (28540026); <b>Blood protein levels (28240269)</b>; <b>Crohn's disease (21102463)</b>; Cervical cancer (28806749); <b>Inflammatory bowel disease (28067908)</b>; <b>Neonatal lupus (20662065)</b>; Psychosis (atypical) (24132900); <b>Ulcerative colitis (28067908)</b></p>	<p><b>positive regulation of I-kappaB kinase/NF-kappaB signaling; positive regulation of NF-kappaB transcription factor activity; toll-like receptor 3 signaling pathway; TRIF-dependent toll-like receptor signaling pathway; death-inducing signaling complex assembly; tumor necrosis factor-mediated signaling pathway; toll-like receptor 4 signaling pathway; cellular response to tumor necrosis factor; positive regulation of defense response; NIK/NF-kappaB signaling; extrinsic apoptotic signaling pathway; cellular response to cytokine stimulus; positive regulation of immune response; apoptotic process; regulation of defense response; I-kappaB kinase/NF-kappaB signaling; immune response; cell surface receptor signaling pathway</b></p>	<p><b>Autoimmune-disease; Immunity</b></p>
WDR72	<p>adolescent idiopathic scoliosis (30019117); blood urea nitrogen measurement (29403010; 31152163); chloride measurement (29403010); chronic kidney disease (29124443; 31152163); chronic kidney disease, serum creatinine measurement (20383146); corpus callosum mid-anterior volume measurement (31530798); <b>erythrocyte count (30595370)</b>; fear of severe pain measurement (28701861); glomerular filtration rate (27588450; 29124443; 29403010; 30604766; 31015462; 31152163; 31451708); glomerular filtration rate, serum creatinine measurement (26831199); hyperuricemia (29124443); kidney stone (31729369); longevity (20834067); macula measurement (30535121); potassium measurement (29403010); renal system measurement (22797727); renal system measurement, blood urea nitrogen measurement (22797727); serum creatinine measurement (20139978; 29124443; 29403010); smoking initiation (30643251); smoking status measurement (30643251); sodium measurement (29403010); <b>type II diabetes mellitus (30297969)</b>; uric acid measurement (29124443); urinary pH measurement (30476138).</p>	<p>amelogenesis; biomineral tissue development; odontogenesis of dentin-containing tooth; positive regulation of enamel mineralization; positive regulation of tooth mineralization; regulation of enamel mineralization; regulation of tooth mineralization</p>	<p>Other</p>

\*Disease/Trait (PUBMEDID) with p-value < 1E-5 as reported in the EBI GWAS catalog

\*\* Biological Process (GO) with p-value < 1E-5 as reported in the STRING