



REVIEW ARTICLE

PERTINENCE DES MARQUEURS MOLECULAIRES SSRs ET SNPs POUR L'ANALYSE DE LA DIVERSITE GENETIQUE DES POPULATIONS NATURELLES: UNE META-ANALYSE

¹ADAMOU IBRAHIM Maman Laouali, ¹BAGUE OUMAROU Zeinabou, ²IDI SAIDOU Sani, ¹ALIO MOUSSA Abdourazak and ¹BAKASSO Yacoubou

¹Département de Biologie de la Faculté des Sciences et Techniques de l'Université Abdou Moumouni de Niamey;
²Université de Diffa.

ARTICLE INFO

Article History:

Received 25th July, 2025
Received in revised form
06th August, 2025
Accepted 19th September, 2025
Published online 30th October, 2025

Keywords:

Genetic Structure; Genetic Diversity;
Molecular Markers; Ssrs, Snps, Statistical
Methods.

*Corresponding author:

ADAMOU IBRAHIM Maman Laouali

ABSTRACT

Genetic diversity is considered as a series of differentiation in the genetic composition of species. This diversity is subject to major environmental constraints. For sustainable management, it is necessary to improve species through knowledge of their genetic structure through well-adapted tools, to face new imperatives. This is how molecular markers such as SSRs and SNPs are used to analyze this structure. The objective of this study is to evaluate the relevance of SSR and SNP markers in the analysis of the genetic structure of natural populations. The use of the PRISMA approach made it possible to identify the articles considered, through keywords launched in search engines such as Google scholar, PubMed and Hal science. A total of 1125 studies were identified, of which twenty-six (26) were included after applying the eligibility, inclusion and exclusion criteria. Among these included studies, sixty-one percent (61%) favored SNP markers against thirty-nine percent (39%) using SSRs. In addition, studies using plant material predominate with a frequency of sixty-eight percent (68%), compared to those using animal material (32%). According to these biological trends, plant studies clearly prefer SNP markers (46.43%), while animal studies use SNPs and SSRs at comparable frequencies (14.29% and 17.86%, respectively). The analysis revealed that various statistical methods or combinations of statistical methods are used to assess the diversity and genetic structure of populations. Statistical indicators used within this meta-analysis revealed highly significant heterogeneity between studies (12=97.3% and P-value < 0.0001), suggesting publication bias due to the methodological limitations intrinsic to small samples, combined with the specificities of genetic markers, type of biological material and data analysis methods.

Copyright©2025, ADAMOU IBRAHIM Maman Laouali et al. 2025. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: ADAMOU IBRAHIM Maman Laouali, BAGUE OUMAROU Zeinabou, IDI SAIDOU Sani, ALIO MOUSSA Abdourazak and BAKASSO Yacoubou. 2025. "Pertinence des marqueurs moléculaires ssrs et snps pour l'analyse de la diversité génétique des populations naturelles: une méta-analyse". *International Journal of Current Research*, 17, (10), 35046-35060.

INTRODUCTION

La génétique des populations est une discipline basée sur la diversité génétique et l'étude de sa répartition au sein des individus, des populations et entre les populations (Becheler, 2013). Cette diversité varie sous l'effet de quatre forces évolutives à savoir la migration, la dérive génétique, la sélection et la mutation (Durand, 2011). Ces forces majeures sont à la base de la structure de la variabilité génétique et de son évolution (Montaigne, 2011). L'étude de cette structure retrace l'évolution de l'espèce et identifie le niveau de variation génétique pour son intégration dans le programme d'amélioration génétique (Nsibi, 2021). La diversité génétique comme composante essentielle de la biodiversité, décrit la variabilité des gènes inter et intra spécifiques et de leur population (Montaigne, 2011). Elle est définie comme la variation dans la composition génétique des espèces (Hvilsomet *et al.*, 2022). La diversité génétique est extrêmement importante, car elle constitue le support de base que peut agir la sélection (Montaigne, 2011). Ainsi, il a été démontré que plus cette diversité est importante dans un groupe d'individu, plus certaines espèces auront la capacité de s'adapter à de nouvelles conditions environnementales (Montaigne, 2011). L'intérêt de la variabilité génétique est qu'elle est non seulement nécessaire à la survie des organismes (Hobanet *et al.*, 2023), mais aussi permet aux espèces de s'adapter aux variations climatiques, aux changements de leur milieu naturel, ainsi qu'à d'autres contraintes, afin de permettre le développement de nouvelles lignées différentes (Hvilsomet *et al.*, 2022, Abdelbaki, 2024). En plus, elle garantit la sécurité alimentaire et le développement durable (Abdelbaki, 2024). La différenciation génétique intra et inter-population appuie les fonctions écologiques et procure des ressources nécessaires et des services fondamentaux pour l'humanité (Hollingsworth *et al.*, 2020). L'intérêt porté à la diversité génétique montre qu'une gestion durable des espèces est nécessaire pour l'évaluation des ressources génétiques (Mauger *et al.*, 2023). Cependant, dans un contexte de variation environnementale actuellement observé, cette diversité assurant pourtant la survie et le potentiel évolutif des populations naturelles est en pleine essor (Montaigne, 2011), ainsi que l'analyse de l'évolution de la structure génétique neutre et adaptative (Vignon *et al.*, 2024). Toutefois, le maintien de la diversité génétique fournie de base importante aux populations pour leur adaptation aux conditions changeantes (Hvilsomet *et al.*, 2022). Il est donc d'une importance cruciale de sauvegarder cette diversité, en vue d'améliorer les espèces pour faire face à de nouveaux impératifs (Hobanet *et al.*,

2023). Afin d'étudier la variabilité ainsi que la structure génétique des populations naturelles, il est important de faire recours à des outils bien adaptés (Mauger *et al.*, 2023). Les marqueurs SSRs et les SNPs développés récemment, sont les plus utilisés pour l'analyse de la diversité et de la structure génétique (Vidal, 2016 ; Mauger *et al.*, 2023). En effet, l'on peut se poser la question de savoir si ces deux (2) marqueurs sont plus pertinents que les autres pour l'étude de la structuration génétique des populations naturelles. C'est ainsi que l'objectif principale de cette étude est d'évaluer la pertinence des marqueurs SSR et SNP dans l'analyse de la structure génétique des populations naturelles. Il s'agit spécifiquement de : (1) déterminer la fréquence d'utilisation des marqueurs SNP et SSR pour l'analyse de la diversité génétique des populations, (2) évaluer l'efficacité et les limites de ces marqueurs (SNP, SSR) dans l'analyse de la diversité génétique, (3) identifier les méthodes statistiques utilisées en analyse de la diversité et de la structure génétique des populations et (4) évaluer les biais de publications liés à l'analyse de la structuration génétique des populations.

METHODOLOGIE

Cette étude a été réalisée suivant les directives PRISMA « Preferred Reporting Items for Systematic Reviews and Meta-Analyses » (Hammaniet *al.*, 2024), une nouvelle version de PRISMA mise en place par Moher *et al.* (2009). Selon Hmanaet *al.* (2020), il s'agit d'une revue systématique qui consiste à trouver de milliers d'articles sur la base des mots clés en fonction d'un certain nombre de critère (stratégies d'inclusion/exclusion et des critères d'éligibilité). Ce programme fait suite à la formulation de la question de recherche et précède la synthèse et l'évaluation de la qualité des articles sélectionnés ainsi que la formulation des recommandations (Sébastien, 2020).

Recherche documentaire: Pour l'identification des articles pertinents à inclure dans cette étude, la recherche littéraire a été réalisée au moyen des moteurs de recherche comme : « Google scholar », « Hal Science » et « PubMed ». Les combinaisons des mots clés suivant, à la fois en français et en anglais ont été utilisées pour la collecte des publications : « *Structure génétique, Diversité génétique, Marqueurs moléculaires, marqueurs SSRs et SNPs* », « *Genetic structure, Genetic diversity, Molecular markers, SSR and SNP markers* ».

Critères d'éligibilité, d'inclusion et d'exclusion : Dans le cadre de cette méta-analyse, ont été éligibles, toutes études : (1) publiées en français ou en anglais, (2) moléculaires incluant les marqueurs SSRs et/ou SNPs, (3) portant sur les espèces animales ou végétales.

Lors de la consultation des bases de données, toutes les publications trouvées en ligne ont été sélectionnées à partir des titres, des résumés disponibles et des textes entiers. Aussi, ne sont incluses que les études ayant été obtenues avec les textes complets et remplissant les critères d'éligibilité. Cependant, ont été éventuellement exclus les articles suivants : les doublons, les titres et les résumés seulement, les titres non concernés, les études sur les traits phénotypiques et les études non éligibles.

Extraction des données: Pour chacune des études retenues, les extraits d'informations suivants ont été enregistrés sur un tableau Excel : le nom du premier auteur, le pays de réalisation de l'étude, l'année de publication de l'étude, le type de matériel biologique, le type de marqueurs utilisé, le nombre de marqueurs utilisé, la précision du résultat, le nombre de groupes génétiques, l'indice de fixation (Fis), l'indice de différenciation génétique (Fst), le taux d'hétérozygotie (He), le nombre d'allèles (Na), le nombre d'individus (N) et le contenu d'information du polymorphisme (PIC).

Analyses des données: Pour la saisie des données collectées ainsi que la réalisation des tableaux, le tableur EXCEL du Microsoft Office 2016 a été utilisé. Le logiciel R (Version 4.2.1) a été utilisé pour la réalisation des graphiques et les analyses statistiques. Ces analyses sont les suivantes:

- L'analyse en composantes principales (ACP), pour décrire les relations entre les variables quantitatives étudiées ainsi que leurs contributions dans la formation des différents axes ;
- La classification hiérarchique ascendante (CHA), pour structurer en groupes les différentes études en fonction des variables observées ;
- La méta-analyse a été réalisée pour évaluer l'hétérogénéité entre les études incluses.

Les indicateurs statistiques utilisés dans cette méta-analyse sont les suivants :

- Les intervalles de confiances (IC) : estime la précision de l'effet global observé. Il indique la plage de valeurs dans laquelle l'effet réel est susceptible de se situer, avec un certain niveau de confiance (95%). Un IC à 95% signifie que si l'étude est répétée plusieurs fois, 95% des intervalles calculés contiendraient l'effet réel (Borensteinet *al.* 2009). Pour un effet moyen $\hat{\theta}$ et son erreur standard SE ($\hat{\theta}$), l'IC à 95% est calculé suivant la formule :

$$IC \text{ à } 95\% = \hat{\theta} \pm 1,96 \times SE(\hat{\theta})$$

Une estimation plus précise de l'effet est montrée par un IC étroit, tandis qu'un IC large indique une incertitude de la précision (Cumpstonet *al.* 2019).

- I-squared (I^2) : mesure le pourcentage de variabilité totale entre les études due à l'hétérogénéité. I^2 ne dépend pas du nombre d'études considéré (Higgins *et al.*, 2002).

$$I^2 = \frac{Q - (k - 1)}{Q} \times 100$$

(Q = Statistique de Cochran, k = Nombre d'études incluses dans la méta-analyse, $(k - 1)$ = Degré de liberté).

Si $I^2 < 0$, I est considéré comme nul (les études sont totalement homogènes) ;

Si $I^2 < 25\%$ cela indique une hétérogénéité faible (les études sont relativement homogènes) ;

Si $25 < I^2 < 75\%$ on a une hétérogénéité modérée ;

Si $I^2 > 75\%$ on parle d'hétérogénéité importante.

• Q de Cochran (Q) : évalue l'hypothèse nulle selon laquelle toutes les études partagent un effet commun (c'est-à-dire qu'il n'y a pas d'hétérogénéité) (Cochran, 1954).

$$Q = \sum w_i (y_i - \bar{y})^2$$

Avec w_i = poids de la $i^{\text{ème}}$ étude, y_i = Effet observé de la $i^{\text{ème}}$ étude, $\bar{y}$ = Effet moyen pondéré).

Une valeur élevée de Q par rapport au degré de liberté ($k - 1$) suggère une hétérogénéité significative.

- P-value : associée à la statistique Q de Cochran (1954), P-value teste l'hypothèse nulle d'absence d'hétérogénéité. Le niveau de significativité a été établi à 0,05.
- **P-value $\leq 0,05$** : rejet de l'hypothèse nulle, indiquant une hétérogénéité significative.
- **P-value $> 0,05$** : pas de preuve suffisante pour rejeter l'hypothèse nulle, suggérant une hétérogénéité non significative.
- Tau-squared (Tau² ou τ^2) : estime la variance réelle entre les effets des études. Il représente la variabilité des effets réels autour de la moyenne globale (DerSimonian et Laird, 1986).

$$\tau^2 = \frac{Q - (k - 1)}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}}$$

(Q = Statistique de Cochran, k = Nombre d'études incluses dans la méta-analyse, $(k - 1)$ = Degré de liberté, w_i = poids de la $i^{\text{ème}}$ étude).

Une valeur élevée de τ^2 indique une grande hétérogénéité entre les études.

- Tau simple (Tau ou τ) : représente la racine carrée de τ^2 . Il présente l'écart-type des effets réels entre les études. Une valeur élevée de τ indique une grande dispersion des effets réels entre les études.
- Si $\tau = 0$: aucune hétérogénéité
- Si $\tau < 0,2$: faible hétérogénéité
- Si $0,2 \leq \tau < 0,5$: hétérogénéité modérée
- Si $\tau > 0,5$: forte hétérogénéité
- H-statistique (H) : évalue l'hétérogénéité globale (Higgins et Thompson (2002)). Il est calculé comme suit :

$$H = \sqrt{\frac{Q}{k - 1}}$$

$H = 1$ et $H < 1,2$ indiquent une absence d'hétérogénéité et $H > 1,5$ montre une hétérogénéité.

RESULTATS

Etudes sélectionnées : Au total, mille cent vingt-cinq (1125) études ont été incluses. Parmi celles-ci, six (6) doublons ont été exclus, trois cent cinquante-huit (358) titres non concernés ont été exclus, six cent quarante-deux (642) présentant le titre et ou abstract seulement exclues, quatre-vingt-douze (92) études qui ont fait cas des traits phénotypiques ont été exclues et une étude non éligible a été exclu, pour enfin retenir vingt-six (26) études dont trois (3) thèses et vingt-trois (23) articles (Figure 1).

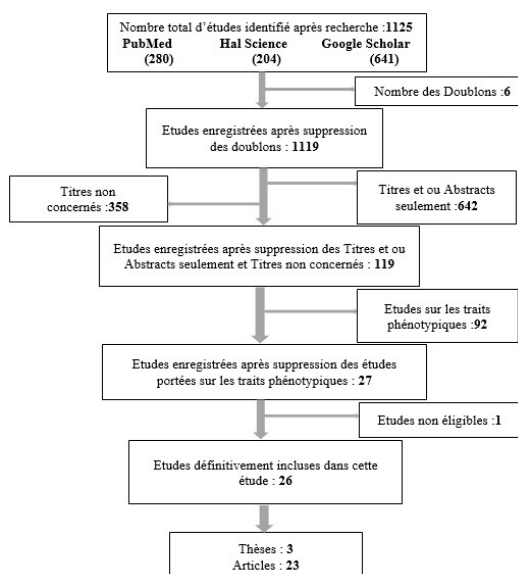


Figure 1. Etudes considérées après la recherche documentaire

Répartition géographique des études : L'analyse de l'usage des marqueurs pour l'évaluation de la diversité génétique montre que la Chine est de loin le pays où l'outil moléculaire est le plus couramment utilisé (46,43%), suivi de loin du Japon avec seulement 7,14% (Figure 2).

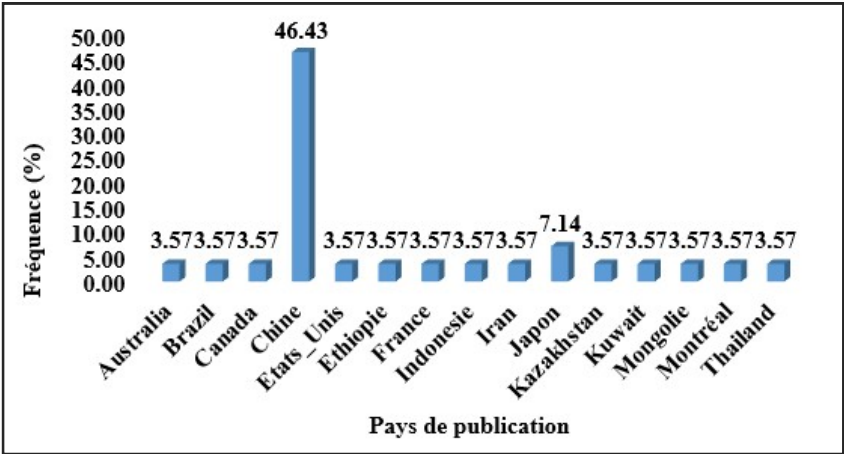


Figure 2. Fréquence de répartition des publications en fonction du pays

Evolution temporelle des études : Les résultats des études des 15 dernières années de 2009 à 2024 (Figure 3). Ainsi, de 2009 à 2017, il n'y a pas eu de variation dans la publication. Des fréquences constantes ont été observées (3,57%). A partir de l'année 2018, il y'a eu une légère évolution (7,14%). Cependant, de 2020 à 2024, une évolution considérable des études a été observée. L'évolution des études considérées est beaucoup plus récente, avec un accroissement significatif de 2020 à 2024.

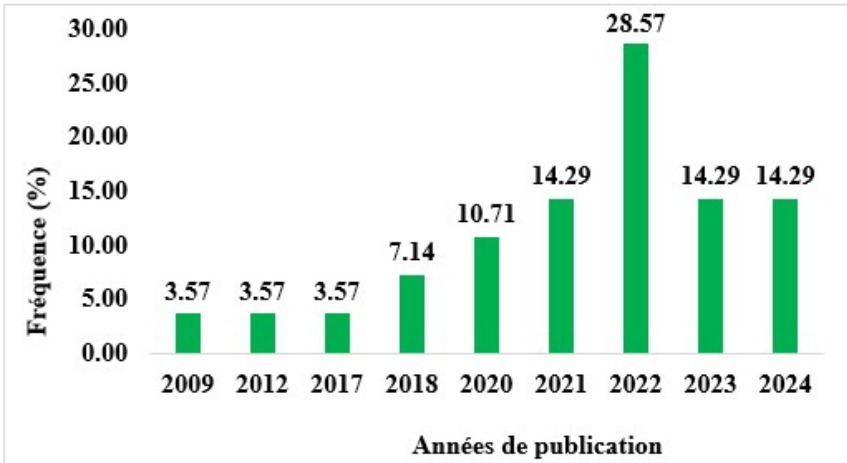


Figure 3. Fréquence de répartition des études en fonction des années

Les résultats de l'étude de l'utilisation des marqueurs SSRs et SNPs dans le temps (2009 à 2024) sont représentés par la figure 4. L'examen de cette figure montre que de 2009 à 2018, les fréquences d'utilisation sont trop faibles (3,57%) et ne varient pas. L'analyse révèle également que les études ont initialement utilisé les marqueurs SSRs de 2009 à 2012 puis les SNPs en 2017 avec des faibles fréquences. A partir de 2018 les deux (2) marqueurs sont utilisés dans les études avec prépondérance des SNPs dont le pic a été atteint en 2022 avec 17,86% d'utilisation des SNPs pour 10,71% des SSRs.

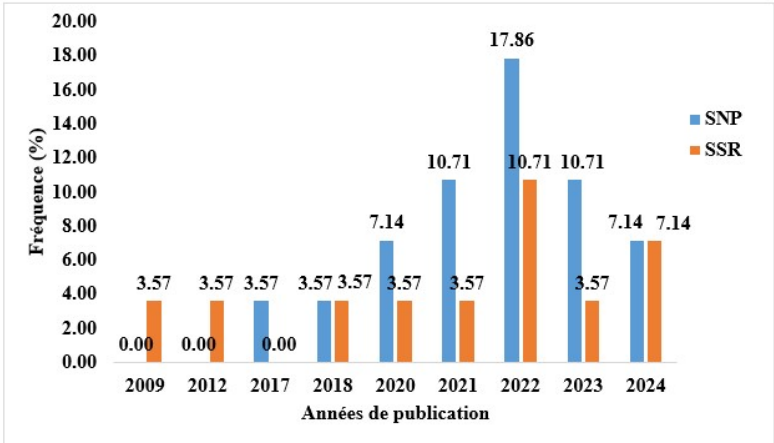


Figure 3. Fréquence de répartition des études en fonction des années

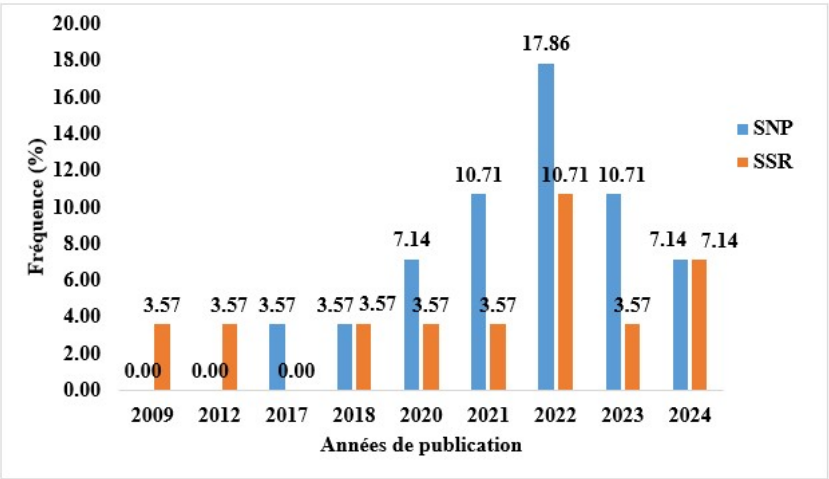


Figure 4. Usage des marqueurs SNPs et SSRs en fonction des années

Répartition des études par type de matériel biologique : La figure 5 présente les fréquences de répartition des études selon les types de matériel biologiques. L’analyse montre que 68% des études ont fait usage du matériel végétal et 32% ont fait cas du matériel animal. Ceci prouve une prédominance des études porté sur le matériel végétal.

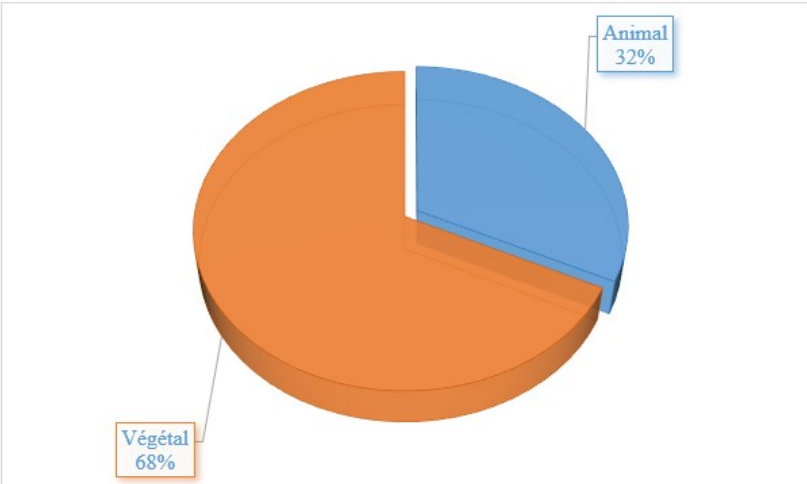


Figure 5. Fréquence de répartition des études selon les types de matériel biologique

Répartition des études par type de marqueur moléculaire : L’analyse de répartition des études considérées en fonction des types de marqueurs moléculaires utilisés montre que 61% des études ont fait recours aux marqueurs SNPs et 39% des études ont utilisées les marqueurs SSRs (Figure 6).

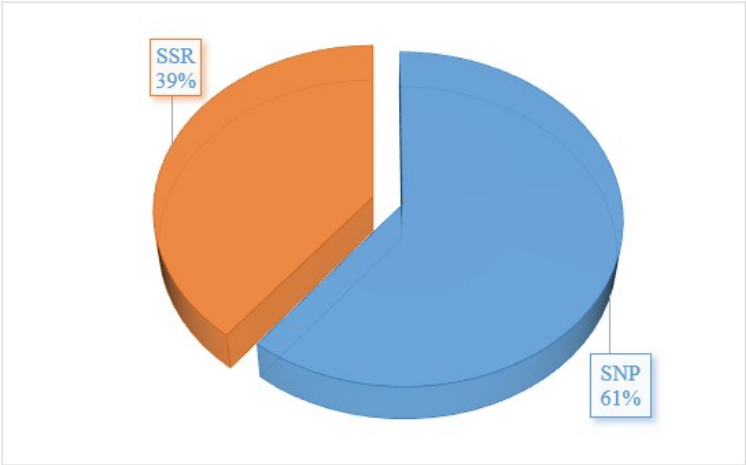


Figure 6. Fréquence de répartition des études selon le type de marqueurs

Interaction entre type de matériel biologique et choix des marqueurs par les études : Les résultats des études sur l'usage des deux marqueurs (SNPs et SSRs) en fonction du type de matériel biologique sont représentés par la figure 7. Parmi les études portant sur le matériel végétal, 46,43% ont fait usage des marqueurs SNPs et 21,43% ont fait recours aux marqueurs SSRs. Les études portant sur le matériel végétal privilégient les marqueurs SNPs. Par contre dans les études portant sur le matériel animal, 17,86% ont fait recours aux marqueurs SSRs et 14,29% ont utilisés les marqueurs SNPs. Les études faisant usage du matériel animal présentent pratiquement les mêmes fréquences d'utilisation de ces deux marqueurs avec une légère supériorité des SSRs par rapport aux SNPs, mais la différence n'est pas statistiquement significative ($P\text{-Value} = 0,424 > 0.05$).

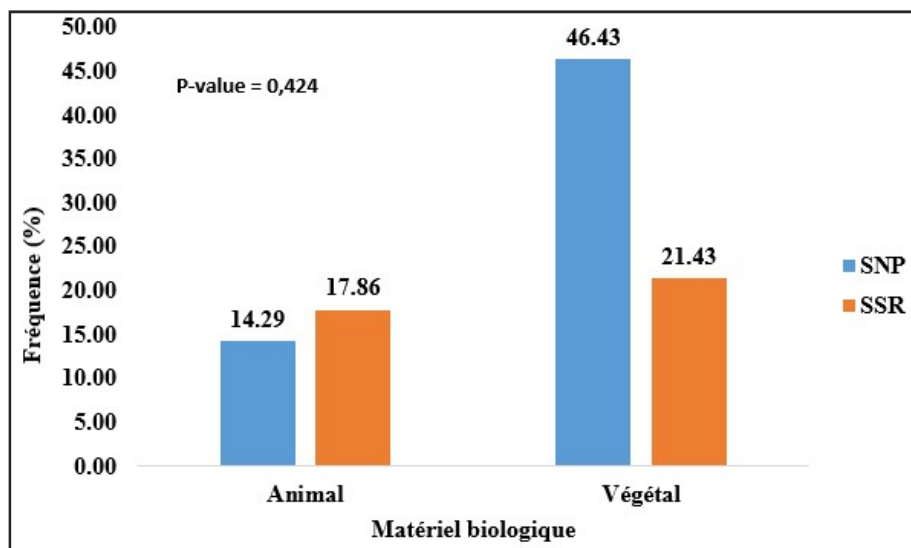


Figure 7. Usage des marqueurs SNPs et SSRs en fonction du type de matériel biologique

Méthodes statistiques utilisées en fonction des études: Les différentes combinaisons de méthodes statistiques utilisées dans les études incluses sont présentées dans le tableau 1. L'analyse a permis d'identifier pour les plus importantes les associations Neighbor-Joining et Structure (32%); Neighbor-Joining, Structure, Analyse Moléculaire de la Variance et Analyse en Composantes Principales (21%) ; et Neighbor-Joining, Structure et Analyse en Composantes principales (21%).

Tableau 1. Méthodes statistiques appliquées dans les études

Méthodes statistiques	N	Fréquence (%)	Fréquence cumulées (%)
NJ et ACP	1	4	4
Nj et AMOVA	1	4	8
NJ et Structure	9	32	40
NJ, Structure et ACP	6	21	61
NJ, Structure et ACoP	2	7	68
NJ, Structure, AMOVA et ACP	6	21	89
NJ, Structure, AMOVA et ACoP	2	7	96
NJ, Structure, AMOVA, et ACP	1	4	100

Légende : ACoP : Analyse en Coordonnées Principales, ACP : Analyse en Composantes Principales, AMOVA : Analyse Moléculaire de la Variance, NJ :Neighbor-Joining, N : nombre d'étude.

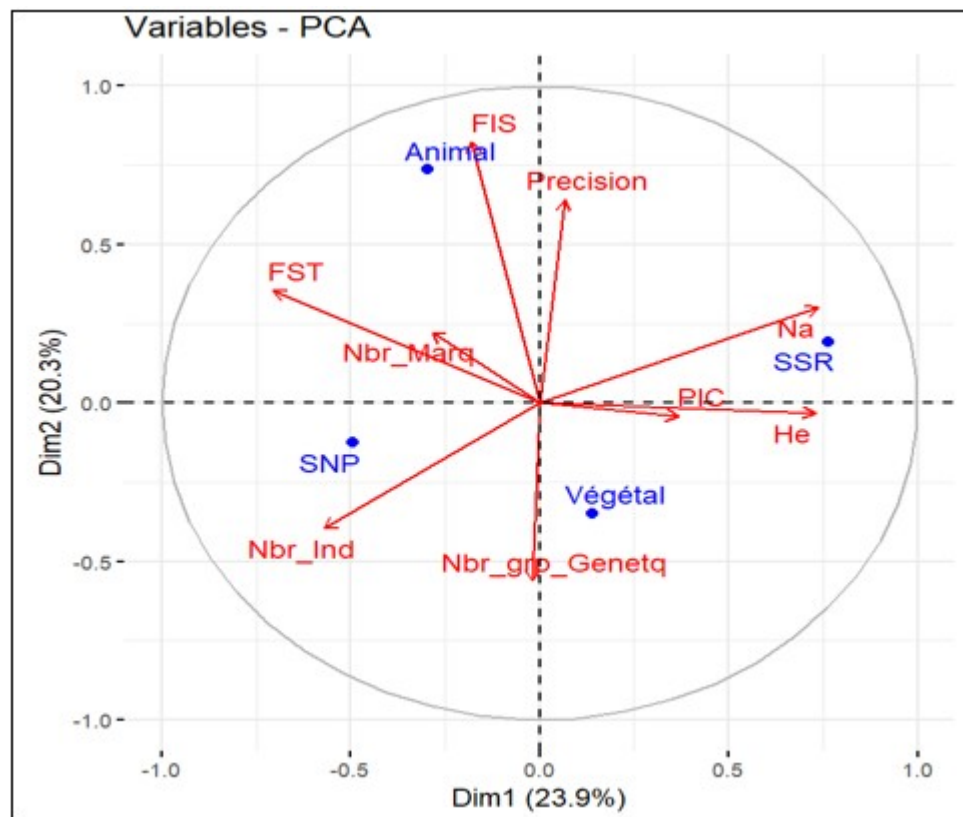
Analyse des relations entre les paramètres énumérés: Les résultats de l'Analyse en Composantes Principales (ACP) ont révélé que les quatre (4)composantes expliquent 70,77% de la variabilité totale avec une valeur propre variant de 1,05 à 2,15 (Tableau 2). Les pourcentages d'inertietotale de chaque axe sont également présentés et varient de 11,76% à 23,92%.

Tableau 2. Valeurs propres et variabilité des axes

Dimensions	Valeur propre	Variance (%)	Variance cumulée (%)
Comp1	2,153	23,920	23,920
Comp2	1,831	20,342	44,262
Comp3	1,326	14,738	59,000
Comp4	1,059	11,768	70,768
Comp5	0,885	9,835	80,602
Comp6	0,626	6,952	87,554
Comp7	0,575	6,389	93,944
Comp8	0,401	4,451	98,395
Comp9	0,144	1,605	100,000

La Figure 8 présente le cercle de corrélation et projection des neuf (9) variables dans le plan formé par les axes 1 et 2 de l'ACP en fonction de leur qualité de représentation dans le plan. L'analyse révèle que les variables des études associées aux longues flèches (Na, He, Fis, Fst,

précision, Nbr_grp_Genetq) sur le plan factoriel présentent les contributions les plus fortes à la formation du cercle de corrélation, reflétant leur poids dans la structuration des données. La projection des variables dans le plan formé par les deux premiers axes de l'ACP, correspond à 44, 26% de la variabilité totale. La dimension 1 (axe 1) est positivement associée aux variables Na et He regroupées au tour des études végétales utilisant les marqueurs SSRs. A l'opposé, Fst et Nbr_Indmontrent une corrélation négative avec cet axe, suggérant que les études avec les SSRs, bien que riche en Na, He, capturent moins de Fst. La dimension 2 (axe 2) se distingue par une corrélation positive avec Fis et précision des résultats se réunissant autour des études animales ayant utilisées les marqueurs SNPs. A l'inverse, leNbr_grp_Genetq est faiblement corrélé négativement avec cet axe.



Légende : SNP : polymorphisme d'un seul nucléotide ; SSR : simple séquence répétée ; Na : nombre d'allèle ; Nbr_Marq : nombre de marqueurs génétique ; Nbr_Ind : nombre d'individus ; Nbr_grp_Genetq : nombre de groupe génétique ; PIC : contenu d'information du polymorphisme ; He : hétérozygotie attendue ; Fis : coefficient de consanguinité ; Fst : indice de différenciation génétique ; Precision : précision du résultat.

Figure 8. Projection des paramètres identifiés dans un plan formé par les dimensions 1 et 2 de l'ACP

La corrélation des axes avec les différentes variables est présentée dans le tableau 3. L'analyse montre que, la dimension 1 (axe 1) associe positivement les variables nombre d'allèle ($r = 0,739$) et l'hétérozygotie attendue ($r = 0,729$) et négativement le nombre d'individus ($r = -0,570$) et l'indice de différenciation génétique ($r = -0,705$). La dimension 2 (axe 2) est positivement corrélée avec la précision des résultats ($r = 0,641$) et la variable indice de consanguinité ($r = 0,823$). Cet axe est négativement corrélé avec la variable nombre de groupe génétique ($r = -0,562$). La dimension 3 (axe 3) est caractérisée par une corrélation positive avec les variables nombre d'individus ($r = 0,557$) et le contenu d'information du polymorphisme ($r = 0,779$). La dimension 4 (axe 4) présente une corrélation positive avec la variable nombre de marqueur génétique ($r = 0,746$).

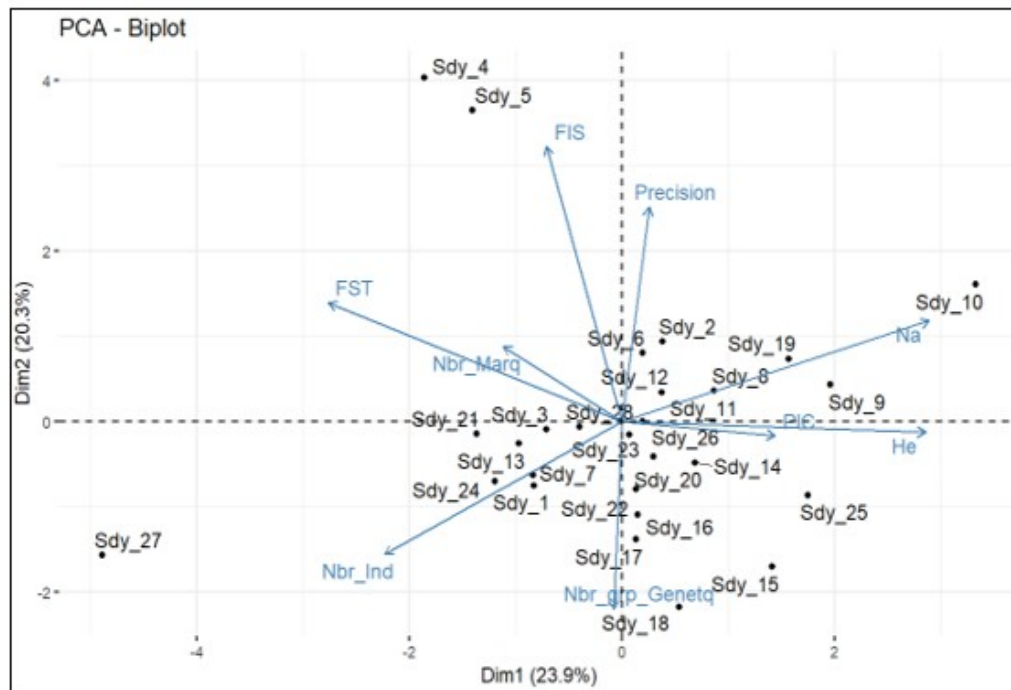
Tableau 3. Corrélations des différentes variables avec les cinq premières composantes

Paramètres	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
Nbr_Marq	-0,285	0,223	0,298	0,746	-0,222
Nbr_Ind	-0,570	-0,396	0,557	-0,144	0,013
Precision	0,065	0,641	-0,121	-0,396	0,365
Nbr_grp_Genetq	-0,020	-0,562	0,071	0,277	0,726
PIC	0,368	-0,042	0,779	-0,211	0,136
FIS	-0,182	0,823	0,210	0,178	0,252
FST	-0,705	0,354	0,301	-0,206	-0,055
Na	0,739	0,302	0,124	0,326	0,125
He	0,729	-0,032	0,389	-0,156	-0,274

Légende : SNP : polymorphisme d'un seul nucléotide ; SSR : simple séquence répétée ; Na : nombre d'allèle ; Nbr_Marq : nombre de marqueurs génétique ; Nbr_Ind : nombre d'individus ; Nbr_grp_Genetq : nombre de groupe génétique ; PIC : contenu d'information du polymorphisme ; He : hétérozygotie attendue ; Fis : coefficient de consanguinité ; Fst : indice de différenciation génétique ; Precision : précision du résultat.

Analyse des relations entre les études et les paramètres : La projection de l'ensemble des études incluses en fonction des variables dans le plan formé par les deux axes (1 et 2) de l'ACP est présentée par la figure 9. La dispersion en nuage des points représente les études. Plusieurs groupes d'études se distinguent par leurs valeurs élevées sur des paramètres spécifiques. Les sdv 4 et 5 se caractérisent par des valeurs élevées de Fis et

une précision analytique accrue. Les sdy 9, 10, 19, et 25 montrent des valeurs élevées de Na et He. Enfin, les sdy 1, 3, 7, 13, 21, 24, 27 et 28 présentent des valeurs élevées de Fst, nombre de marqueurs génétiques et nombre d'individus.



Légende : Sdy : étude ; SNP : polymorphisme d'un seul nucléotide ; SSR : simple séquence répétée ; Na : nombre d'allèle ; Nbr_Marq : nombre de marqueurs génétique ; Nbr_Ind : nombre d'individus ; Nbr_grp_Genetq : nombre de groupe génétique ; PIC : contenu d'information du polymorphisme ; He : hétérozygotie attendue ; Fis : coefficient de consanguinité ; Fst : indice de différenciation génétique ; Précision : précision du résultat.

Figure 9. Projection des études en fonction des paramètres identifiés dans un plan formé par les dimensions 1 et 2 de l'ACP.

Analyse de la structuration de la variabilité des études : Pour mieux voir la répartition des études dans la figure 14, la classification ascendante hiérarchique (CAH) a été réalisé (Figure 10). L'analyse CHA est une méthode d'analyse statistique utilisée afin de regrouper de manière plus claire des éléments similaires en fonction des différentes caractéristiques. Il y'a donc une homogénéité intra-groupe et une hétérogénéité inter-groupe. Cette classification a produit un dendrogramme de troncature gradué et a reparti les études en trois (3) groupes selon la méthode de Ward. Le premier groupe (G1) est constitué de quatre (4) études (sdy 4, 5, 27 et 21) présentées en rouge sur le dendrogramme et sur le cluster plot. Le deuxième groupe (G2) coloré en vert, regroupe six (6) études (sdy 18, 10, 15, 25, 19 et 9) sur le dendrogramme et dix-huit (18) études (sdy 1, 3, 6, 23, 12, 11, 16, 13, 7, 17, 14, 24, 28, 26, 2, 8, 22, et 20) sur le cluster plot. Et enfin, le troisième groupe (G3) coloré en bleu, est formé de dix-huit (18) études (sdy 1, 3, 6, 23, 12, 11, 16, 13, 7, 17, 14, 24, 28, 26, 2, 8, 22, et 20) sur le dendrogramme et six (6) études (sdy 18, 10, 15, 25, 19 et 9) sur le cluster plot. La projection des différents groupes dans le plan factoriel formé par les axes canoniques (1 et 2) montre que les études sont bien structurées en trois (3) groupes.

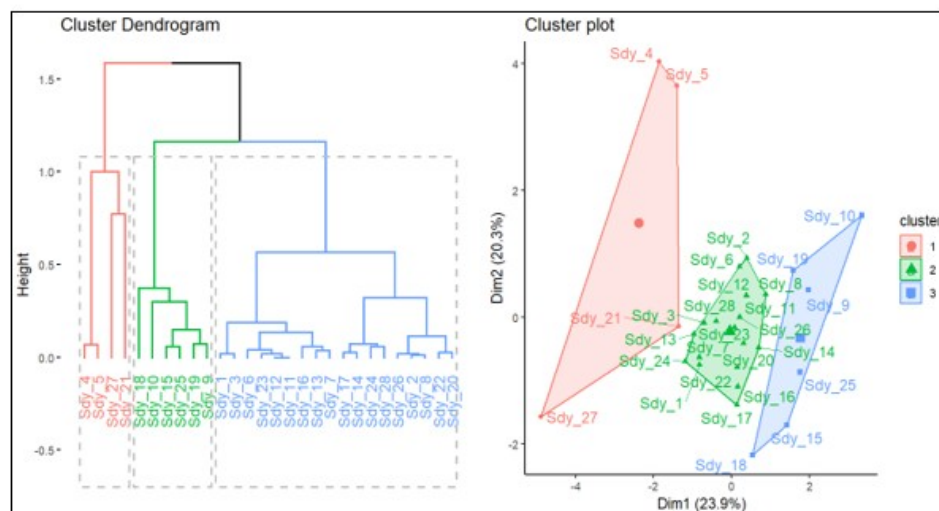


Figure 10. Dendrogramme issu de la classification des études et représentation des groupes dans le plan formé par les premiers axes de l'ACP

Analyse d'hétérogénéité des études et recherche de biais : Le forest plot (diagramme en forêt aléatoire) est une représentation graphique utilisée dans les revues systématiques ou méta-analyses pour comparer les résultats issus de différentes études abordant une même question. Le forest plot des études globales est présent en quatre colonnes (Figure 11) dont : (1) auteurs, dates, pays ; (2) groupes génétiques obtenus pour chaque étude ; (3) nombre d'échantillon considéré ; (4) et les différentes proportions avec intervalles de confiance à 95%. Cette figure fait

ressortir une hétérogénéité très hautement significative entre les études ($I^2 = 97,3\%$ et $P\text{-value} < 0,0001$). Certaines études se sont montrées homogènes et possèdent un intervalle de confiance très faible, tandis que d'autres se sont différenciées ou écartées et disposent d'un intervalle de confiance un peu plus élevé. En effet, au sein d'un même pays, certaines études présentent une homogénéité entre elles, tandis que d'autres se sont éloignées, c'est le cas de la Chine.

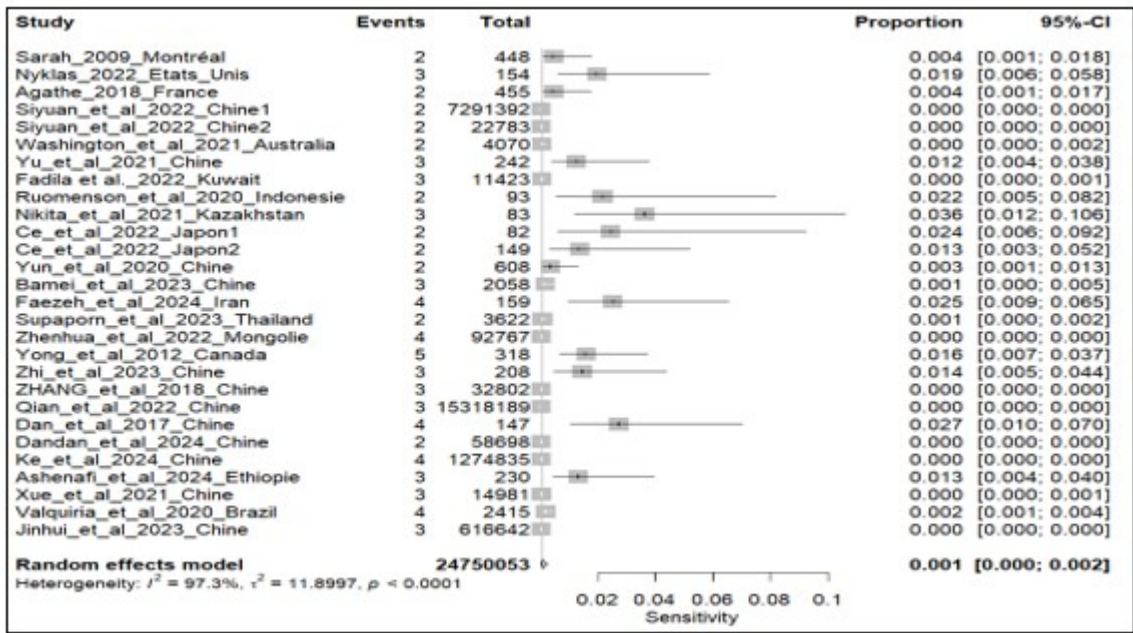


Figure 11. Forest plot de toutes les études incluses.

Le funnel plot (graphique en entonnoir) est une représentation en nuage de points permettant de détecter visuellement la présence d'un biais de publication ou d'hétérogénéité, en fonction de la distribution des erreurs standard des études incluses individuellement dans une méta-analyse (Figure 12). Sur cette figure, le funnel plot est très asymétrique, ce qui explique la présence d'un biais de publication. Toutes les études possèdent des erreurs standards élevées (c'est-à-dire qu'elles sont réparties sur l'axe horizontal). Un grand nombre d'études est surreprésenté en dehors du rectangle.

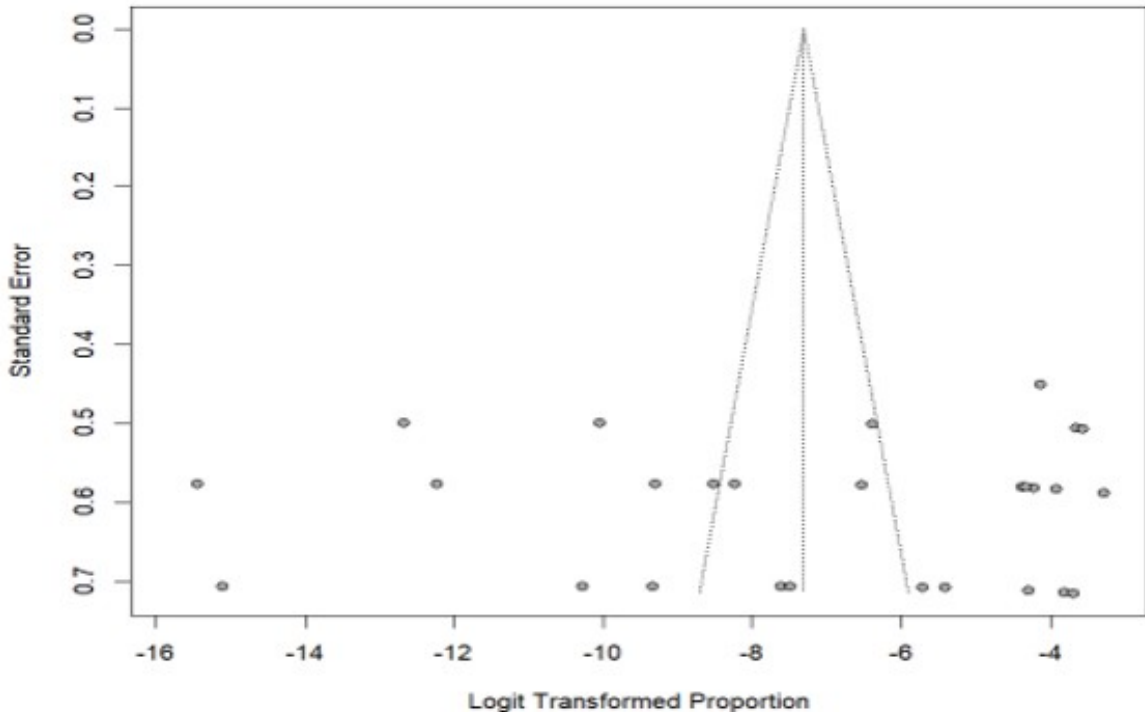


Figure 12. Funnel Plot des de toutes les études incluses

Le forest plot des études utilisant les marqueurs SNPs (Figure 13) montre une hétérogénéité importante entres les études ($I^2 = 97,3\%$ et $P\text{-Value} < 0,0001$). Cette hétérogénéité est créée par deux études de la Chine et une étude réalisée au Japon. Ces études se sont différenciées et écartées de la ligne verticale.

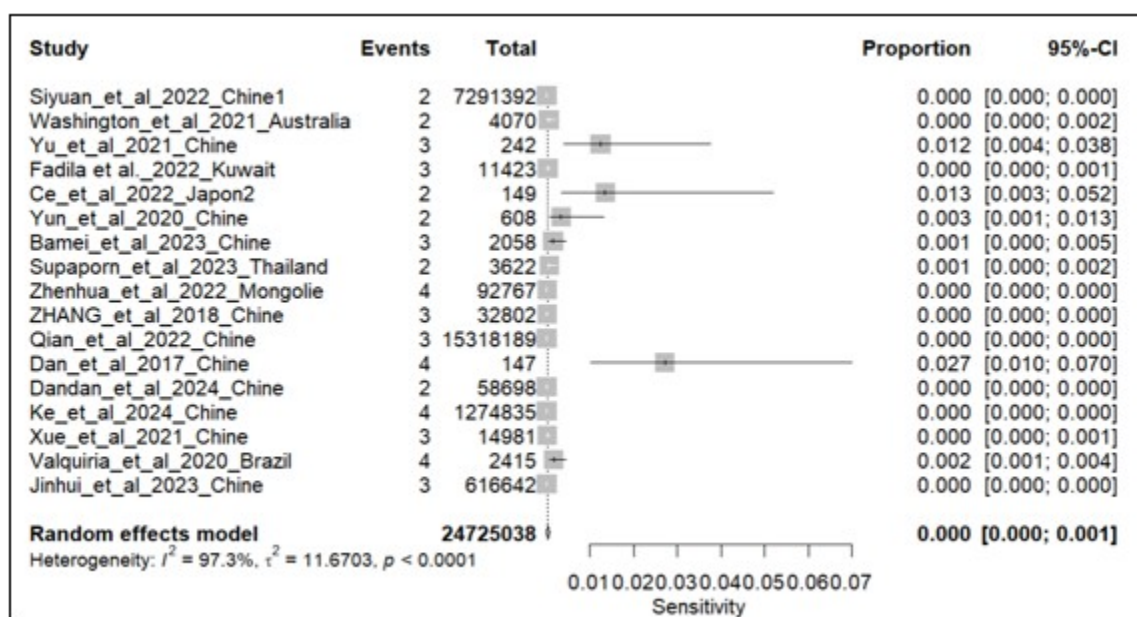


Figure 13. Forest plot des études portées sur les marqueurs SNPs

Le Funnel Plot des études portées sur les marqueurs SNPs (Figure 14) présente une très forte asymétrie. La présence de cette dernière suppose un biais de publication. Les études présentées possèdent des erreurs standards élevées, c'est-à-dire situées sur l'axe horizontal et sont majoritairement distribuées hors du rectangle.

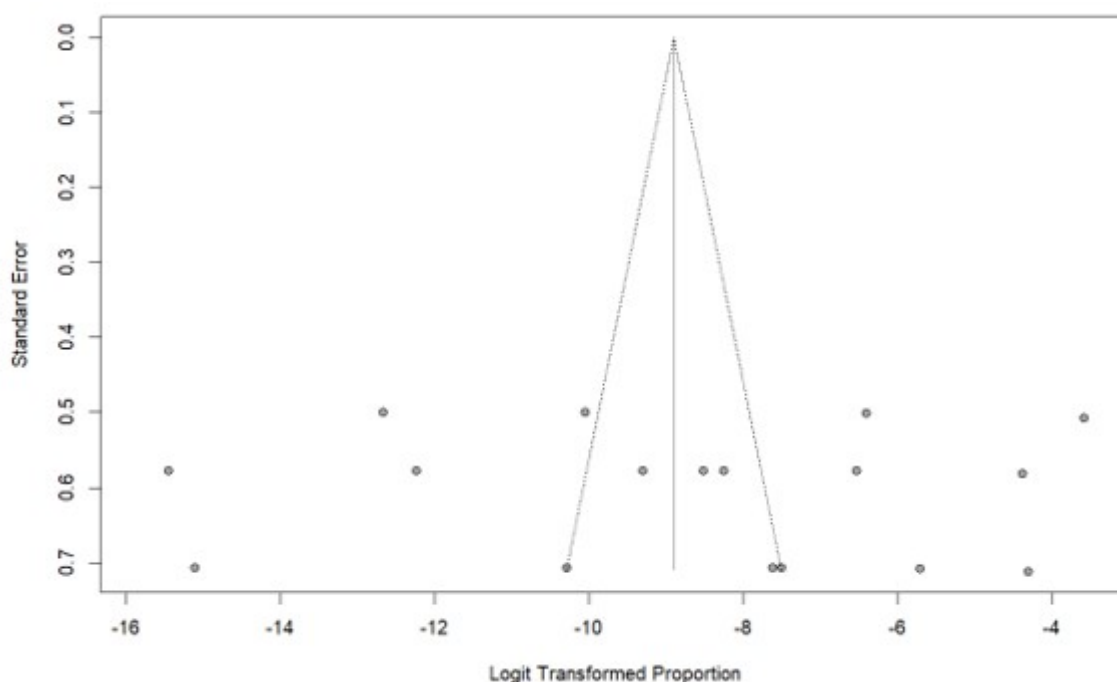


Figure 14. Funnel Plot des études portées sur les marqueurs SNPs

Le forest plot des études faisant recours aux marqueurs SSRs (Figure 15) présente une grande hétérogénéité ($I^2 = 83,7\%$ et $P\text{-Value} < 0,0001$). Cette hétérogénéité est moindre par rapport aux études réalisées sur les SNPs. Des pays comme la Kazakhstan et l'Iran sont un peu plus éloignés des autres pays, avec des intervalles de confiance un peu supérieurs des autres (0,036 et 0,025 respectivement). Le funnel plot des études faisant recours aux marqueurs SSRs (Figure 16), présente moins d'asymétrie. Mais il suggère également un biais de publication. Toutes les études possèdent des erreurs standards relativement élevées (c'est-à-dire une dispersion vers la base du rectangle, sur l'axe horizontal). Seulement deux études se sont écartées du rectangle et sont en majorité représentées au côté droit. Cependant, une seule étude est représentée du côté gauche. Les indicateurs couramment utilisés dans les méta-analyses pour évaluer l'hétérogénéité entre les études incluses sont résumés dans le tableau 6. Ces indices sont calculés pour les études utilisant les deux (2) marqueurs (SNPs et SSRs), ainsi que pour toutes les études incluses. L'analyse globale montre une hétérogénéité très hautement significative avec : (1) τ^2 important ; (2) $I^2 > 75\%$; (3) $H > 1,5$; $Q > k-1$ et $P\text{-Value} < 0,0001$. Les valeurs élevées du Tau, indique une dispersion importante des effets réels entre les études. Cependant, pour les études portant sur les marqueurs SSRs, les valeurs de ces indices sont un peu moins élevées que les autres (Tableau 4).

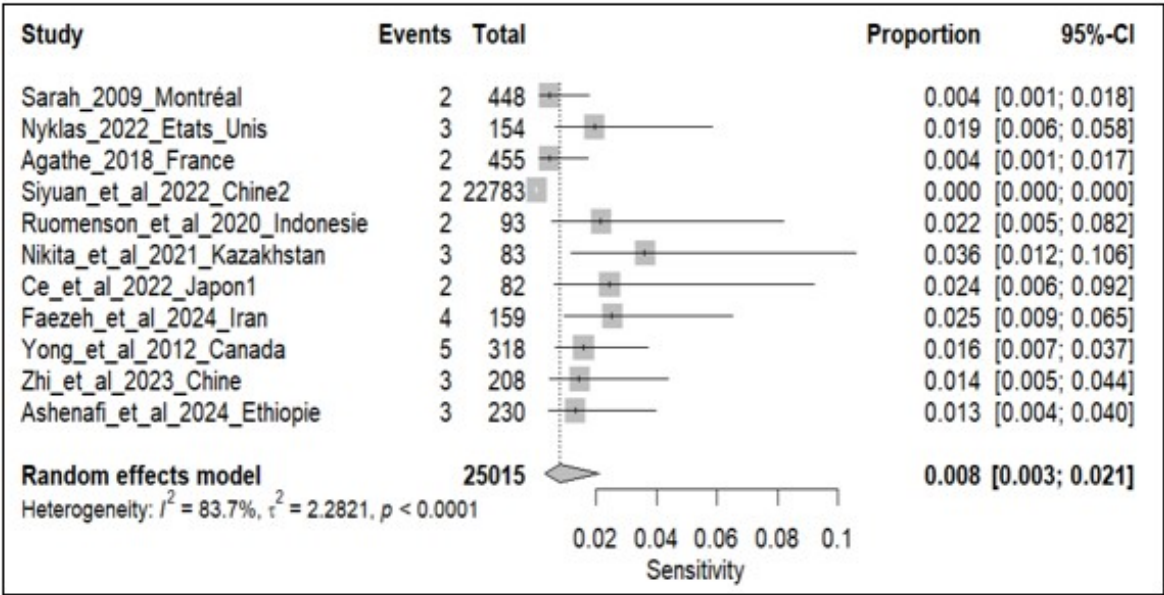


Figure 15. Forest Plot des études portées sur les marqueurs SSRs.

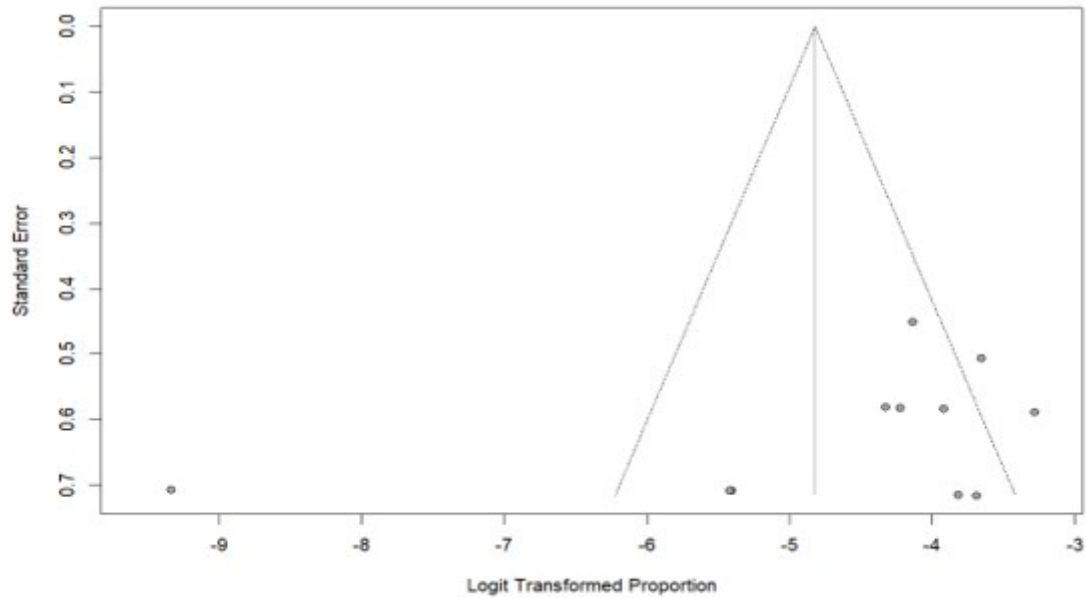


Figure 16 : Funnel plot des études portées sur les marqueurs SSRs.

Tableau 4. Paramètres d’analyse de l’hétérogénéité entre les publications considérées

Type de matériel	Nombre d'études	Tau²	Tau	I²	H	Q	P-value
Toutes les études	28	11,89	3,44	97,3%	6,05	986,98	< 0,0001
SNPs	17	11,67	3,41	97,3%	6,06	587,22	< 0,0001
SSRs	11	2,28	1,51	83,7%	2,48	61,35	< 0,0001

Tableau 5. Synthèse comparative de l’efficacité et des limites des marqueurs SSR et SNP basée sur les résultats de cette étude

Critère	Marqueurs SSR	Marqueurs SNP
Diversité génétique	Haute efficacité	Limité par nature bi allélique
Polymorphisme	Elevé	Réduit
Différenciation génétique	Faible résolution	Excellente résolution
Structure des populations	Performance faible	Performance élevée
Nombre de marqueurs	Limité	Très élevé
Avantages clés	- Détection des allèles rares, - idéale pour les petits échantillons, - informativité par locus élevée	Génotypage haut-débit
Limites principales	- Standardisation difficile, - Nombre limité de loci informatifs	Sous-estimation de la diversité

Efficacité et limites des marqueurs SSR et SNP: Les marqueurs SSR (ou microsatellites), basés sur la variation du nombre de répétitions de courtes séquences d'ADN et les marqueurs SNP, reflétant une variation d'un seul nucléotide, sont deux piliers de la génomique. Bien que complémentaires, leurs efficacités et leurs limites diffèrent significativement selon le contexte d'étude, comme le résume le tableau 5.

DISCUSSION

L'étude moléculaire consiste à évaluer la variabilité entre génotype au niveau du matériel génétique (Ibrahim, 2020). Les marqueurs microsatellites (SSRs) et le polymorphisme d'un seul nucléotide (SNP) développés récemment, sont devenus de marqueurs génétiques importants pour l'analyse de la diversité génétique (Ibrahim, 2020). L'objectif général de cette étude est d'évaluer la pertinence des marqueurs SSR et SNP dans l'analyse de la structure génétique des populations naturelles. Au total, mille cent vingt-cinq (1125) études ont été obtenues pour cette méta-analyse, parmi lesquelles vingt-six (26) ont été incluses. La Chine constitue le pays où l'outil moléculaire est le plus couramment utilisé, probablement grâce à un fort soutien gouvernemental, une infrastructure de recherche avancée dans le domaine des biotechnologies (séquençage génétique et analyse des données génomiques) et une politique active de publication scientifique. Ces avancées technologiques permettent une utilisation plus efficace et plus large des marqueurs moléculaires comme les SNPs et les SSRs. Ricroch (2020), souligne l'avancement chinois dans le dépôt de brevets biotechnologiques et l'application industrielle de ces innovations. Cependant, le faible taux des études réalisées en dehors de la Chine peut être dû par des déficits d'expertise technique, des contraintes financières ou peu d'intérêts par certains professionnels des études moléculaires. L'analyse des résultats révèle que les études de diversité et de génotypage faisant recours aux marqueurs SNPs sont privilégiées par rapport à celles utilisant les SSRs. Ceci s'explique par la standardisation technologique (automatisation et séquençage à haut débit facilitant l'agrégation de données multicentriques), la stabilité (variation ponctuelle avec un taux de mutation faible), la haute densité génomique et la compatibilité des SNPs avec les outils bio-informatiques modernes pour des besoins en haute résolution nécessitant des données massives. Les études utilisent moins les marqueurs SSRs, car ces derniers sont moins adaptés pour les analyses comparatives à grande échelle. MacArthur *et al.* (2017), souligne que 90% des études d'association récentes utilisent des SNPs. Les études ont initialement utilisé les marqueurs SSRs, grâce à leur polymorphisme élevé, idéal pour les études de diversité génétique et de parenté à petite échelle. La transition vers les SNPs s'explique par l'avènement du séquençage haut débit (NGS) permettant leur identification massive à moindre coût. Les SNPs sont plus standardisés et adaptés aux puces de génotypage. L'évolution des études utilisant simultanément les deux marqueurs (SNPs et SSRs), avec partout une prépondérance des SNPs dans les années récentes est due à des innovations technologiques, des besoins scientifiques émergents et des contraintes pratiques liées aux méthodes classiques. Les SNPs restent dominant grâce à leur couverture génomique étendue et leur intégration dans les outils bio-informatiques modernes. Les études sur le matériel végétal sont les plus rencontrées dans cette méta-analyse par rapport à celles réalisées sur le matériel animal, pour des raisons liées à leur diversité, leur pertinence écologique et leur utilité pour étudier les mécanismes évolutifs. Les méta-analyses en biodiversité restent dominées par les plantes, occultant d'autres lignées évolutives (Antonelli *et al.*, 2023). En plus, les plantes structurent les écosystèmes et réagissent rapidement aux perturbations environnementales. Par contre, la faible fréquence des études utilisant le matériel animal peut être dû à leur échantillonnage complexe et coûteux due à la forte mobilité des espèces animales. Les études végétales privilégient les marqueurs SNPs, par rapport aux SSRs. Les plantes dotées d'une forte diversité, leur analyse nécessite des marqueurs nombreux et stables comme les SNPs pour couvrir leur génome, alors que les SSRs peuvent être saturés rapidement du fait de cette forte diversité compliquant ainsi les analyses à grande échelle. Par contre, les études faisant usage du matériel animal présentent presque les mêmes fréquences d'utilisation de ces deux marqueurs, mais avec une légère supériorité des SSRs par rapport aux SNPs. Ceci pourrait être dû à leur complémentarité adaptée aux défis spécifiques des études en milieu naturel. Cette complémentarité reflète un équilibre lié à des traditions méthodologiques, des contraintes pratiques et des innovations technologiques, notamment dans un contexte de crise de la biodiversité où chaque outil a son rôle.

L'analyse des études incluses a révélé les combinaisons des méthodes statistiques les plus importantes pour l'analyse de la structure et de la variabilité génétiques dont Neighbor-Joining et Structure; Neighbor-Joining, Structure, Analyse Moléculaire de la Variance et Analyse en Composantes Principales; Neighbor-Joining, Structure et Analyse en Composantes Principales. Ces méthodes se combinent pour leur complémentarité, afin de permettre une analyse multidimensionnelle efficace. Elles possèdent la capacité d'explorer, d'analyser et d'interpréter des données alléliques ou génotypiques complexes générés par les études moléculaires (Lamara, 2010). La structuration de la variabilité entre les études incluses par les analyses multivariées (ACP, CHA) a fait ressortir trois (3) groupes distincts. Le premier groupe (G1) regroupe les études animales utilisant les marqueurs SNPs. Ce groupe d'études se caractérise par des valeurs élevées des paramètres de structure populationnelle tel que le F_{is} (Wright, 1949), le F_{st} (Wright, 1978) ainsi que la précision des résultats et le nombre d'individus. Des valeurs élevées du F_{is} et F_{st} dans ce groupe d'étude s'expliquent respectivement par une consanguinité et une différenciation génétique entre sous-populations. L'indice de consanguinité $F_{is} = (H_e - H_o)/H_e$ varie théoriquement de -1 (en cas d'excès d'hétérozygotes par rapport à l'attente sous Hardy-Weinberg) à 1 (si tous les individus sont homozygotes, traduisant ainsi une consanguinité maximale). L'indice de différenciation génétique $F_{st} = (H_t - H_s)/H_t$ varie de 0 (en cas d'absence de différenciation entre sous-populations) à 1 (cas de différenciation complète entre sous-populations). Ces indices sont souvent employés dans des contextes de sauvegarde ou de maintien des ressources animales (Bhatia *et al.*, 2013). Quant à la précision, elle renvoie à la robustesse des analyses de clustering (structure), influencée par le nombre d'individus et de marqueurs (Pritchard *et al.*, 2000). Le deuxième groupe (G2) est un groupe moyennement intermédiaire du groupe (G1) et (G3), réunissant d'une part des éléments du groupe G1 et d'autre part des éléments du groupe G3. Cette transition pourrait probablement être liée à l'utilisation combinée des marqueurs SSRs et SNPs par les études intégrant à la fois des organismes animaux et végétaux. Pour une vision holistique de la biodiversité, certaines études soulignent la nécessité d'intégrer différents marqueurs (Fuentes-Pardo *et al.*, 2017). En fin, le troisième groupe (G3) se distingue par les études végétales faisant recours aux marqueurs SSRs. Ce groupe d'étude montre des valeurs élevées des indices de diversité génétique telle que le nombre d'allèle, l'hétérozygotie attendue de chaque locus (Nei, 1987) le contenu d'information du polymorphisme. En effet, ces indicateurs de diversité génétique élevés dans une population suggèrent respectivement une haute discrimination des loci, un nombre important d'allèles et une large diversité génétique au sein des populations. Les marqueurs SSRs sont largement utilisés en génétique végétale en raison de leur capacité à capturer une diversité génétique fine (Kalia *et al.*, 2011). L'organisation des études en trois (3) groupes s'explique par les différences dans les types de marqueurs génétiques utilisés (SNPs et SSRs), les organismes étudiés (animaux et végétaux), ainsi que les paramètres statistiques associés. Les indicateurs d'hétérogénéité ont montré une différenciation hautement significative entre les études incluses. Dans la présentation du forest plot, les vastes études (nombre d'échantillon élevé), reflètent des intervalles de confiance étroits et une homogénéité relative. Par contre, les études plus petites (faible nombre d'échantillon) présentent des intervalles de confiance large, créant une grande variation entre les études et une contribution majeure à l'asymétrie du Funnel plot. Cette asymétrie, conforme aux observations de Fanelli (2018), signale un biais de publication. L'hétérogénéité et l'asymétrie observées dans cette étude reflètent donc des limites méthodologiques liées aux petits échantillons, aux spécificités des marqueurs génétiques, aux types de matériel biologique et aux méthodes statistiques utilisées.

CONCLUSION

Cette méta-analyse a révélé que les études faisant usage des marqueurs SNPs sont privilégiées par rapport à celles faisant recours aux marqueurs SSRs dans les études de diversité et structure génétique. En plus, selon les tendances biologiques, les études sur les végétaux ont été plus réalisées à l'aide des marqueurs SNPs, tandis que celles portant sur les animaux sont réalisées à l'aide des SNPs et des SSRs. Ces choix s'expliquent par les avantages distincts des deux technologies. Les SNPs, bien que moins polymorphes, se révèlent adaptés aux analyses de structure génétique à grande échelle, alors que les SSRs, plus variables, permettent une résolution fine de la diversité génétique. En plus, les combinaisons des méthodes statistiques les plus importantes pour l'analyse de la structure et de la variabilité génétiques ont été révélées. Les résultats ont également montré une hétérogénéité hautement significative entre les études, suggérant un biais de publication. L'analyse dérivée des paramètres d'hétérogénéité a montré que l'hétérogénéité se situe plus dans les études impliquant les marqueurs SNPs. L'hétérogénéité et l'asymétrie observées dans cette méta-analyse reflètent donc des limites méthodologiques intrinsèques aux petits échantillons, combinées aux spécificités des marqueurs génétiques, du type de matériel biologique et aux méthodes statistiques utilisées. Cependant, ni les SNPs, ni les SSRs ne sont universellement supérieurs. Leur pertinence dépend du contexte biologique et spatio-temporel de l'étude. La complémentarité de ces marqueurs, couplée à une interprétation critique des sources d'hétérogénéité, reste la clé pour déchiffrer la structuration génétique des populations naturelles.

RECOMMANDATIONS

Pour une étude de la diversité et de la structure génétique, nous recommandons aux chercheurs les points suivants:

- Optimiser le choix des marqueurs selon l'organisme et les objectifs.
- Standardiser les protocoles méthodologiques : définir des critères minima pour la taille des échantillons, le nombre de loci analysés, ainsi que les méthodes statistiques.
- Atténuer les biais de publication en encourageant la publication des données via les financements des recherches dans les pays sous-représentés pour limiter la sélection arbitraire des résultats.
- Renforcer les collaborations inter-pays, en créant des consortiums internationaux pour partager les infrastructures techniques (séquenceurs haut débit), en uniformisant les protocoles d'échantillonnage et en ciblant les régions sous-représentées.

Ces propositions visent à transformer les défis identifiés (hétérogénéité, biais) en leviers pour une recherche plus inclusive et rigoureuse.

Conflit d'intérêt : Les auteurs déclarent qu'il n'y a pas de conflit d'intérêt.

Remerciement : Les auteurs remercient le département de biologie de l'université Abdou Moumouni de Niamey pour avoir facilité la conduite de cette étude.

REFERENCES

- Abdelbaki, B. (2024). Contribution à la caractérisation génétique de la population unicolore locale : Étude de la variabilité génétique et groupes phénotypiques. Thèse de doctorat, Université de Mostaganem (Algérie). 184p.
- Ahn, E., & Kang, H. (2018). Introduction to systematic review and meta-analysis. *Korean Journal of Anesthesiology*, 71(2), 103–112.
- Akriche, I., & Mechri, R. (2019). Utilisation des marqueurs morphologiques et biochimiques pour l'analyse de la diversité génétique de l'espèce d'avoine tétraploïde *Avena magna*. 65p.
- Anthony, B. (2020). Étude des ressources génétiques du noyer en vue de la mise en œuvre d'une sélection assistée par marqueurs. Biologie végétale. Université de Bordeaux (France). Français. Tel-03118714f. 329p.
- Antonelli, A., Dhanjal-Adams, KL, & Silvestro, D. (2023). Integrating machine learning, remote sensing and citizen science to create an early warning system for biodiversity. *Plants, People, Planet*, 51A (3), 307-316.
- Becheler, R. (2013). Diversité génétique d'espèces structurantes en environnement marin : influence sur la réponse démographique des populations aux perturbations anthropiques. Thèse de doctorat, Université de Bretagne occidentale-Brest (Bretagne). Français. NNT: 2013BRES0055. Tel-01065808. 179p.
- Bérard, C., Tanguay, C., & Bussièrès, J. F. (2014). Revue de la littérature reproductible. *Annales de l'Unité de recherche en pratique pharmaceutique*, 1-9.
- Ben Ayed, R. (2017). Applications et limites des marqueurs moléculaires pour l'identification des variétés d'olivier. *Agriculture*, 8(1), 49-54.
- Bhandari, H.R., Bhanu, A. N., Srivastava, K., Singh, M. N., & Shreya, H. A. (2017). Assessment of genetic diversity in crop plants-an overview. *Adv. Plants Agric. Res*, 7 (3), 279-286.
- Bhatia, G., Patterson, N., Sankararaman, S., & Price, A. L. (2013). Estimating and interpreting FST : the impact of rare variants. *Genome research*, 23 (9), 1514-1521.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. 421p
- Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, 10(1), 101-129.
- Cumpston, M., Li, T., Page, M. J., Chandler, J., Welch, V. A., Higgins, J. P., & Thomas, J. (2019). Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *Cochrane Database of Systematic Reviews*, 10 (2019), Article ED000142.
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials*, 7 (3), 177-188.
- Diaby, A. (2024). Revue de littérature et Méta-analyse de la Séroprévalence du Virus de l'Hépatite B au Mali de 1990 à 2021. Thèse de doctorat, Université des Sciences, des Techniques et des Technologies de Bamako (Mali). 90p.
- Durand, E. (2011). Etude des bases (épi) génétiques de l'adaptation dans une expérience de sélection divergente pour la précocité de floraison chez le maïs. Thèse de doctorat, Université de Paris Sud-Paris XI (France). Français. NNT: 2011PA112078. Tel-00607712. 228p.

- Fallah, M. (2020). La construction d'une carte génétique consensus à haute densité chez le soja basée sur des marqueurs SNP dérivés du génotypage par séquençage (GBS). 60p
- Fanelli, G., & Serretti, A. (2019). The influence of the serotonin transporter gene 5-HTTLPR polymorphism on suicidal behaviors : a meta-analysis. *Prog Neuropsychopharmacol Biol Psychiatry*, 88, 375-387.
- Fuentes-Pardo, A.P, Ruzzante, D.E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Mol Ecol* 26(20) : 5369-5406.
- Gerber, S., Mariette, S., & Nous, C. (2022). Heurs et malheurs des marqueurs moléculaires. Hal - 03945257. 9p.
- Gupta P.K., Rustgi S., and Mir, R.R. (2001). Array-based high-throughput DNA markers for crop improvement. *Heredity* 101: 5-18.
- Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, 555(7695), 175-182.
- Halewood, M., López Noriega, I., Ellis, D., Roa, C., Rouard, M., & Sackville-Hamilton, N. R. (2017). Implications potentielles de l'utilisation de l'information génétique numérique sur les ressources génétiques sur les trois objectifs de la Convention sur la diversité biologique. Document soumis par le CGIAR au Secrétariat de la Convention sur la diversité biologique. 40p.
- HAMMANI, S., & ANGADE, K., (2024). La Gouvernance Territoriale : Un Pilier Déterminant du Développement territorial- Une Analyse à travers l'Approche de la Proximité : revue systématique de la littérature. *Alternatives Managériales Economiques*, 6(2), 288-307.
- Higgins, J.P., & Thompson, S.G., (2002). Quantifying heterogeneity in meta-analysis. *Stat Med* 2002, 21 (11), 1539-1558.
- HMAMA, Z., & ALAOUI, M. (2020). Entrepreneuriat social et innovation sociale, quelle distinction ? : Une revue de littérature sur la base de la méthode PRISMA. *Revue Internationale du Chercheur*, 1(2). 136p.
- Hoban, S., da Silva, J.M., Mastretta-Yanes, A., Grueber, C.E., Heuertz, M., Hunter, M.E. & Laikre, L. (2023). Monitoring status and trends in genetic diversity for Convention on Biological Diversity : an ongoing assessment of genetic indicators in nine countries. *Conservation Letters*, 16 (3), e12953.
- Hollingsworth, P.M., O'Brien, D., Ennos, R.A., Ahrends, A., Ballingall, K.T., Brooker, R.W., Burke, T., Cavers, S., Dawson, I.K., Elston, D.A., Kerr, J., Marshall, D.F., Neaves, L., Pakeman, R.J., Trivedi, C., Wall, E., Wright, F., Yahr, R., Bean, C., Blake, D., Campbell, R., Comont, R., Finger, A., Fraser, K., Genney, D., Hall, J., Hannah, A., Jehle, R., Jones, S., Kohn, D., Llewellyn, M., Lurz, P., Macdonald, I., McIntosh, J., Mitchell, R., O'Dell, J., Page, S., Pemberton, J., Pérez-Espona, S., Piernney, S., Sime, I., Thompson, D., Ogden, R. (2020). Scotland's Biodiversity Progress to 2020 Aichi Targets: Aichi Target 13– Genetic Diversity Maintained – Supplementary Report 2020. Inverness, Scottish Natural Heritage. 33p.
- Hvilsom, C., Segelbacher, G., Ekblom, R., Fischer, M.C., Laikre, L., Leus, K., O'Brien, D., Shaw, R. & Sork, V. (2022). Selecting species and populations for monitoring of genetic diversity. Publication of IUCN. 47p.
- Ibrahim, A. (2020). Analyse des facteurs socio-environnementaux de production et de la structure génétique des populations du palmier dattier (*Phoenix dactylifera L.*) du Niger. Thèse de doctorat, Université Abdou Moumouni de Niamey (Niger). Biostatistique et génétique des populations. 150p.
- Jay F., Francois O. & Blum M.G.B. (2011). Predictions of Native American Population Structure Using Linguistic Covariates in a Hidden Regression Framework. *PLoS ONE*, 6(1): 16227.
- Kalia, R.K., Rai, M.K., Kalia, S., Singh, R., & Dhawan, A.K. (2011). Microsatellite markers: an overview of recent progress in plants. *Euphytica*, 177 (3), 309-334.
- KHACHTIB, Y. (2023). Caractérisation génétique et moléculaire du pommier au Maroc et de son ravageur le puceron cendré. 150p.
- Kpatenon, M. J. (2022). Diversité génétique et histoire évolutive de deux espèces de palmier utiles (*Borassus aethiopicum Mart.* et *Raphia vinifera P. Beauv.*) en Afrique de l'Ouest et Centrale: implications des changements climatiques et impacts de l'homme (Doctoral dissertation, Université de Montpellier; Université d'Abomey-Calavi (Bénin)). 196p.
- Lamara, M. (2010). Analyse comparative de la diversité génétique et de la structure des populations chez l'orge (*Hordeum vulgare L.*) à l'aide de marqueurs SSR, DArT et du pedigree. 62p.
- Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol. Ecol. Resour.* 19 (3) : 639–647.
- MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., & Parkinson, H. (2017). The new NHGRI-EBI catalog of published genome-wide association studies (catalogue GWAS). *Nucleic Acids Res*, 45 (D1), D896-D901.
- Mauger, S., Baud, A., Le Corguillé, G., Tanguy, G., Legeay, E., Creis, E., & Destombe, C. (2023). Genetic resources of macroalgae: Development of an efficient method using microsatellite markers in non-model organisms. *Algal Research*, 75, 103251.
- Mateo, S. (2020). Procédure pour conduire avec succès une revue de littérature selon la méthode PRISMA. *Kinésithérapie, la Revue*, 20(226), 29-37.
- Mehenni I., Talbi R., (2023). Création d'une base de données pour les Marqueurs moléculaires liés à la résistance aux maladies fongiques chez le blé. Mémoire présenté en vue de l'obtention du diplôme de Master à l'Université Frères Mentouri Constantine (Algérie). Bio-informatique. 71p.
- Michael W. G., (2011). Neighbor Joining, in the Yeasts (Fifth Edition), Volume 1. 40p.
- Mir, R., Shafi, S., et Zargar, S. (2023). Chapter 2 - Genome mapping, in: Mir R, Shafi S, et Zargar S, Principles of Genomics and Proteomics, Elsevier, 29-46, <https://doi.org/10.1016/B978-0-323-99045-5.00007-0>.
- Mittal, B., Chaturvedi, P. et Tulsyan S. (2013). Restriction Fragment Length Polymorphism, in: Maloy S, Hughes K, Brenner's Encyclopedia of Genetics (Second Edition), Academic Press, 190-193, <https://doi.org/10.1016/B978-0-12-374984-0.01314-0>.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of internal medicine*, 151 (4), 264-269.
- Montaigne, W. (2011). Diversité génétique et adaptation au milieu chez les arbres forestiers tropicaux : Etude chez le genre *Vitellaria* (Myristicaceae). Thèse de doctorat, Université des Antilles et de la Guyane (UAG) (France). 178p.
- Nsibi, M. (2021). Contribution des stratégies de sélection génomique et phénotypique aux programmes d'amélioration génétique de l'abricotier (*Prunus armeniaca L.*) pour quelques traits d'intérêt. Thèse de doctorat, Université de Montpellier SupAgro (France). Français. NNT : 2021NSA M0021. Tel-04067940. 214p.
- Pritchard J.K., Stephens M. & Donnelly P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155 (2): 945-959.
- Rasmussen, H.B. (2012). Restriction Fragment Length Polymorphism Analysis of PCR-Amplified Fragment (PCR-RFLP) and Gel Electrophoresis – Valuable Tools for Genotyping and Genetic Fingerprinting. In Techopen. 334p.
- Ricroch, A. (2020). Les biotechnologies en Chine : investissement stratégique et massif dans l'édition du génome. *Monde chinois*, (1), 54-69.
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, 4(4), 406-425.

- Sari, D., Sari, H., Ikten, C., & Tokar, C. (2023). Genome-wide discovery of di-nucleotide SSR markers based on whole genome re-sequencing data of *Cicer arietinum*L. and *Cicer reticulatum*Ladiz. Scientific Reports, 13(1), 10351.
- Schönberg, IJ (1935). Remarks to Maurice Frechet's article ``sur la définition axiomatique d'une classe d'espace distances vectoriellement applicable sur l'espace de Hilbert. Annales de mathématiques, 36 (3), 724-732.
- Simaro, S. (2019). Contribution à l'étude de la structuration génétique et la capacité vectorielle des populations de *Simuliumdamnosum*sl à l'aide de marqueurs microsatellites dans trois faciès épidémiologiques de l'onchocercose en Côte d'Ivoire. Thèse de doctorat, Université Jean Lorougnon Guédé (Côte d'Ivoire). Français. NNT: tel-03739863.157p.
- Singh, P., Mondal, S. et Singh, R.(2021). Chapter 1 - Introduction, in: Mondal S, et Singh R (éditeurs), Advances in Animal Genomics, Academic Press, 1-12, <https://doi.org/10.1016/B978-0-12-820595-2.00001-1>.
- Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method.101 (30), 11030-11035.
- Vidal, M. (2016). Optimization of maritime pine breeding strategies using molecular markers.129p.
- Vignon V., Jean-François A., Anne D., (2024). Étude génétique de l'Agrion de Mercure à l'échelle de l'agglomération Strasbourgeoise et de son autoroute de contournement. Université de Lille (France) – UMR CNRS 8198 Evo-Evo-Paléo / Bureau d'études O.G.E. Dossier technique. P17.
- Wright, S. (1949). La structure génétique des populations. Annales d'eugénisme, 15 (1), 323-354.
- Wright S., (1978). - Evolution and genetics of populations, Vol. 4: Variability within and among natural populations, University of Chicago (USA).
- Yang, W., Kang, X., Yang, Q., Lin, Y. et Fang, M. 2013. Review on the development of genotyping methods for assessing farm animal diversity, Journal of Animal Science and Biotechnology 2013, 4, 1-6.
