

Analyses bio-informatiques de la biologie des microARN : état de l'art, limitations, questions ouvertes

Hervé Seitz

IGH (UMR 9002, CNRS et univ. de Montpellier)

February 1, 2024

Ce diaporama est disponible à :

<http://www.igh.cnrs.fr/equip/Seitz/MOD2024.pdf>

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

miR-656 :

5' AAUAUUUAUCAGUCAACCUCU 3'

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

miR-656 :

5' AAUAUUUAUACAGUCAACCUCU 3'

- ▶ Petite taille, non-poly-adenylés : indétectables en RNA-Seq classique (→ Small RNA-Seq).

Particularités des microARN

miR-656 :

5' AAUAUUUAUACAGUCAACCUCU 3'

- ▶ Petite taille, non-poly-adenylés : indétectables en RNA-Seq classique (→ Small RNA-Seq).
- ▶ Portent un 5' phosphate, 3'-OH ($\neq$ la plupart des produits de dégradation).

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

miR-656 :

5' AAUAUUUAUACAGUCAACCUCU 3'

- ▶ Petite taille, non-poly-adenylés : indétectables en RNA-Seq classique (→ Small RNA-Seq).
- ▶ Portent un 5' phosphate, 3'-OH ($\neq$ la plupart des produits de dégradation).
- ▶ Maturés par des RNases III à partir de précurseurs en tige-boucle ($\neq$ la plupart des produits de dégradation).

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

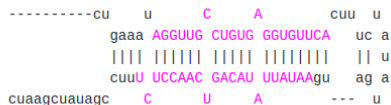
Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

Structure



(structure tirée de [miRBase](#))

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

miR-656 :

5' AAUAUUUAUACAGUCAACCUCU 3'

- ▶ Petite taille, non-poly-adenylés : indétectables en RNA-Seq classique (→ Small RNA-Seq).
- ▶ Portent un 5' phosphate, 3'-OH ($\neq$ la plupart des produits de dégradation).
- ▶ Maturés par des RNases III à partir de précurseurs en tige-boucle ($\neq$ la plupart des produits de dégradation).

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

miR-656 : 5' AAUUAUUAACAGUCAACCUCU 3'
|||||

ARNm ALG13 humain (NM_001324291.2) : 3'...AACAUUUUAUUAUAAUACAUAUAAUUAACUAAAUUAGUG... 5'

- ▶ Petite taille, non-poly-adiénylés : indétectables en RNA-Seq classique (→ Small RNA-Seq).
- ▶ Portent un 5' phosphate, 3'-OH ($\neq$ la plupart des produits de dégradation).
- ▶ Maturés par des RNases III à partir de précurseurs en tige-boucle ($\neq$ la plupart des produits de dégradation).
- ▶ Principal déterminant de reconnaissance des cibles : complémentarité à la « graine ».

Particularités des microARN

Le Small RNA-Seq : spécifique des ARN à 5' phosphate, 3'-OH.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des microARN

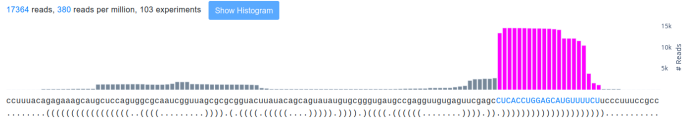
Le Small RNA-Seq : spécifique des ARN à 5' phosphate, 3'-OH.



(profils tirés de [miRBase](#))

Particularités des microARN

Le Small RNA-Seq : spécifique des ARN à 5' phosphate, 3'-OH.



(profils tirés de [miRBase](#))

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des microARN

Répression des cibles

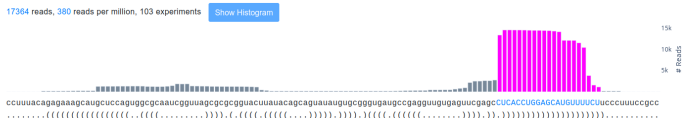
Identification à haut débit des cibles de microARN

Conclusion

Acknowledgements

Particularités des microARN

Le Small RNA-Seq : spécifique des ARN à 5' phosphate, 3'-OH.



(profils tirés de [miRBase](#))

→ Importance du profil de couverture en Small RNA-Seq.

Particularités des microARN

Le Small RNA-Seq : spécifique des ARN à 5' phosphate, 3'-OH.



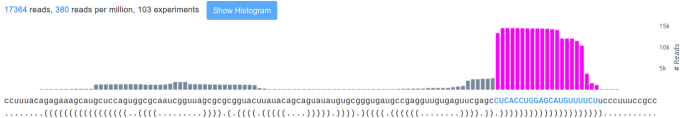
(profils tirés de [miRBase](#))

→ Importance du profil de couverture en Small RNA-Seq.

Pas de moyennage des biais de ligation (Jayaprakash *et al.*, 2011 ; Kim *et al.*, 2019)

Particularités des microARN

Le Small RNA-Seq : spécifique des ARN à 5' phosphate, 3'-OH.



(profils tirés de [miRBase](#))

→ Importance du profil de couverture en Small RNA-Seq.

Pas de moyennage des biais de ligation ([Jayaprakash et al., 2011](#) ; [Kim et al., 2019](#)) → adaptateurs flanqués de poly(N), ...

Répression des cibles

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible.

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

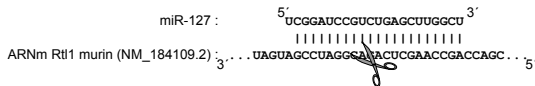
Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible.



Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible.

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Sinon : un effet double :

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Sinon : un effet double :

- ▶ dégradation exonucléolytique (moins efficace) ;

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Sinon : un effet double :

- ▶ dégradation exonucléolytique (moins efficace) ;
- ▶ répression traductionnelle des ARN restants.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Sinon : un effet double :

- ▶ dégradation exonucléolytique (moins efficace) ;
- ▶ répression traductionnelle des ARN restants.

Contributions relatives à la répression : $\approx$ 85% dégradation, 15% répression traductionnelle ([Guo et al., 2010](#)).

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Sinon : un effet double :

- ▶ dégradation exonucléolytique (moins efficace) ;
- ▶ répression traductionnelle des ARN restants.

Contributions relatives à la répression : $\approx$ 85% dégradation, 15% répression traductionnelle ([Guo et al., 2010](#)).

Le Small RNA-Seq est si sensible qu'il détecte des miARN trop peu abondants pour réellement réprimer des cibles ([Mullokandov et al., 2012](#)).

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Répression des cibles

Complémentarité étendue ($\approx$ nt 2–14 parfaitement appariés) : clivage de l'ARN-cible. Rare chez les Bilatériens !

Sinon : un effet double :

- ▶ dégradation exonucléolytique (moins efficace) ;
- ▶ répression traductionnelle des ARN restants.

Contributions relatives à la répression : $\approx$ 85% dégradation, 15% répression traductionnelle ([Guo et al., 2010](#)).

Le Small RNA-Seq est si sensible qu'il détecte des miARN trop peu abondants pour réellement réprimer des cibles ([Mullokandov et al., 2012](#)).

—> dans la littérature, énormément d'assignations fonctionnelles aberrantes.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

- Transcriptomique et protéomiques différentielles.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

- ▶ Transcriptomique et protéomiques différentielles.
- ▶ Cibles directes : immuno-précipitation du complexe effecteur.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

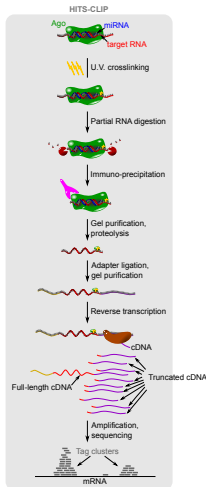
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

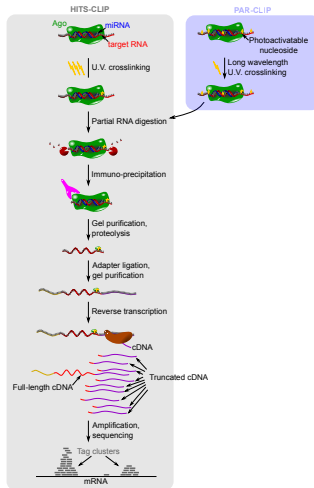
Données
supplémentaires

Identification à haut débit des cibles de microARN



(tiré de Mockly et
Seitz, 2019)

Identification à haut débit des cibles de microARN



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

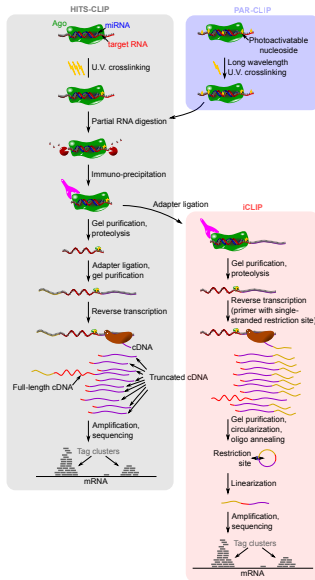
Conclusion

Acknowledgements

Données
supplémentaires

(tiré de Mockly et
Seitz, 2019)

Identification à haut débit des cibles de microARN



(tiré de Mockly et
Seitz, 2019)

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

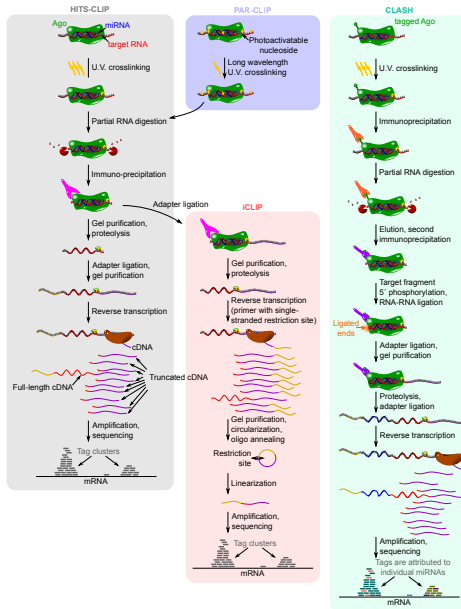
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

(tiré de Mockly et
Seitz, 2019)

Identification à haut débit des cibles de microARN

- ▶ Transcriptomique et protéomiques différentielles.
- ▶ Cibles directes : immuno-précipitation du complexe effecteur.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

- ▶ Transcriptomique et protéomiques différentielles.
- ▶ Cibles directes : immuno-précipitation du complexe effecteur.

Problème conceptuel : observer une répression (ou une interaction physique) ne garantit pas sa fonctionnalité.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

- ▶ Transcriptomique et protéomiques différentielles.
- ▶ Cibles directes : immuno-précipitation du complexe effecteur.

Problème conceptuel : observer une répression (ou une interaction physique) ne garantit pas sa fonctionnalité.

→ Intérêt des prédictions informatiques.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

- ▶ Transcriptomique et protéomiques différentielles.
- ▶ Cibles directes : immuno-précipitation du complexe effecteur.

Problème conceptuel : observer une répression (ou une interaction physique) ne garantit pas sa fonctionnalité.

→ Intérêt des prédictions informatiques.

- ▶ Recherche de *seed matches* dans les 3' UTR
- ▶ Conservées dans l'évolution

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Identification à haut débit des cibles de microARN

- ▶ Transcriptomique et protéomiques différentielles.
- ▶ Cibles directes : immuno-précipitation du complexe effecteur.

Problème conceptuel : observer une répression (ou une interaction physique) ne garantit pas sa fonctionnalité.

→ Intérêt des prédictions informatiques.

- ▶ Recherche de *seed matches* dans les 3' UTR
- ▶ Conservées dans l'évolution

Probablement le meilleur prédicteur : [TargetScan](#).

The screenshot shows the TargetScanHuman website. At the top, the logo "TargetScanHuman" is displayed with the text "Predicting microRNA targets in mammals" and "Release 6.0 September 2011". Below the logo, there are links: "Go to TargetScanHouse", "Go to TargetScanWorm", "Go to TargetScanFly", and "Go to TargetScanFish". The main section is titled "Search for predicted microRNA targets in mammals". It contains a search form with the following fields: 1. "Select a species" (a dropdown menu with "Human" selected). 2. "Enter a human gene symbol (e.g. 'TP53') or an Ensembl gene (ENSG00000148966) or transcript (NM017000403041) ID". 3. "Do one of the following:" with three radio buttons: "Select a broadly conserved microRNA family" (selected), "Select a conserved microRNA family", and "Select a poorly conserved but confidently annotated microRNA family". There are also links to "Select another miRBase accession".

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Limitations de la génomique comparative

Limitations de la génomique comparative

Conservation du *seed match* indépendante du miARN.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Limitations de la génomique comparative

Conservation du *seed match* indépendante du miARN.

≈ 1500 familles de miARN, pour $4^6=4096$ hexamères possibles.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

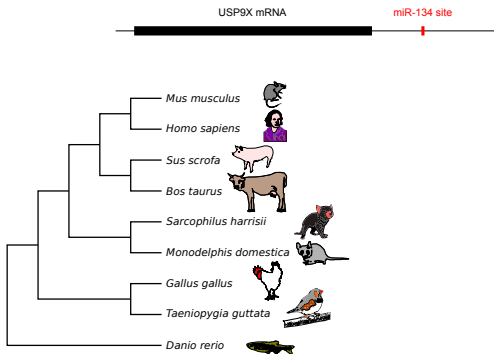
Acknowledgements

Données
supplémentaires

Limitations de la génomique comparative

Conservation du *seed match* indépendante du miARN.

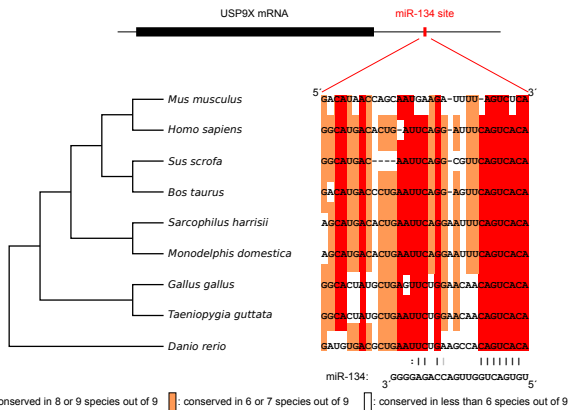
≈ 1500 familles de miARN, pour $4^6=4096$ hexamères possibles.



Limitations de la génomique comparative

Conservation du *seed match* indépendante du miARN.

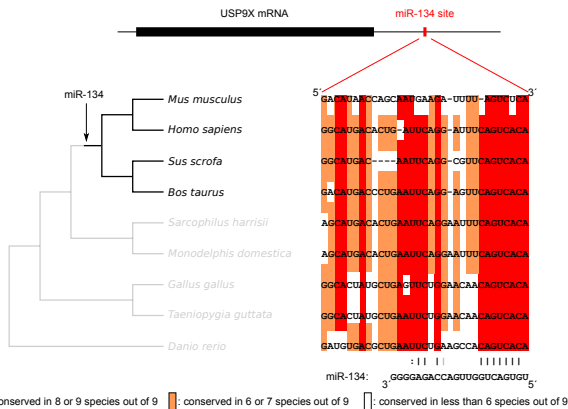
≈ 1500 familles de miARN, pour $4^6=4096$ hexamères possibles.



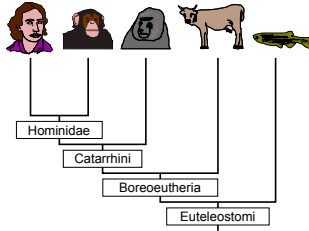
Limitations de la génomique comparative

Conservation du *seed match* indépendante du miARN.

≈ 1500 familles de miARN, pour $4^6=4096$ hexamères possibles.



Limitations de la génomique comparative



(tiré de [Pinzón et al., 2017](#))

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des
microARN

Répression des
cibles

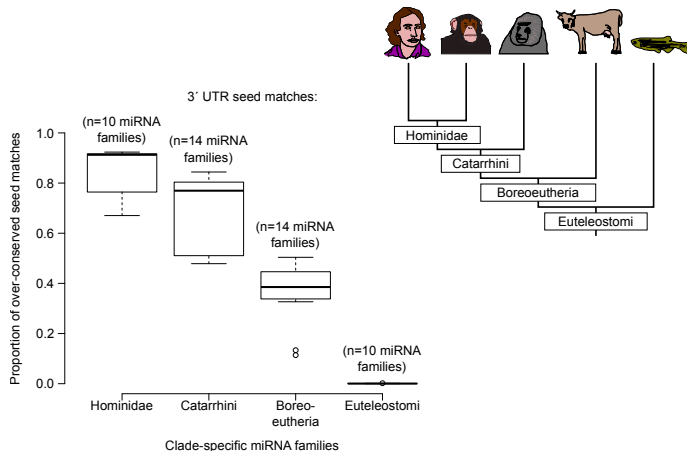
Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

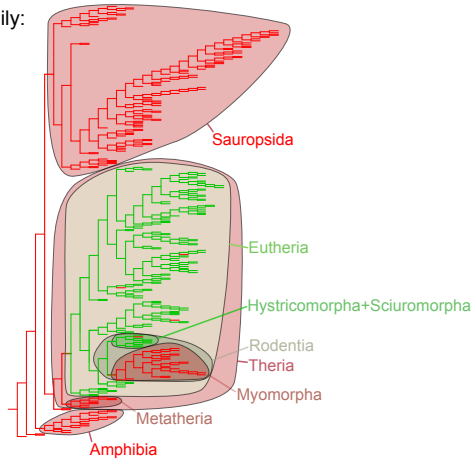
Données
supplémentaires



(tiré de [Pinzón et al., 2017](#))

Limitations de la génomique comparative

miR-656 family:



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

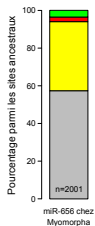
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Limitations de la génomique comparative



- Évolution accélérée dans le clade-frère
- Évolution accélérée dans le clade de perte
- Conservé également dans le clade de perte et le clade-frère
- Pas significativement conservé dans le superclade

Limitations de la génomique comparative

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

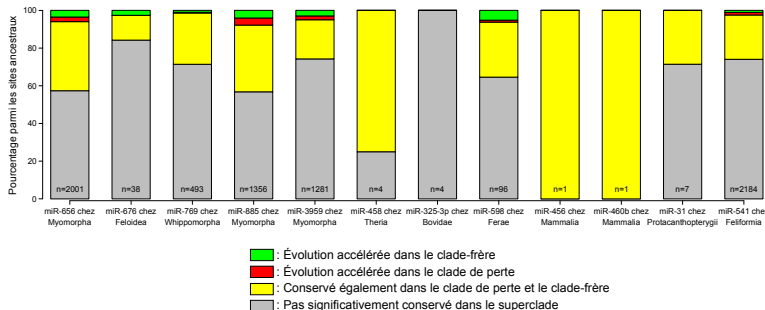
Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

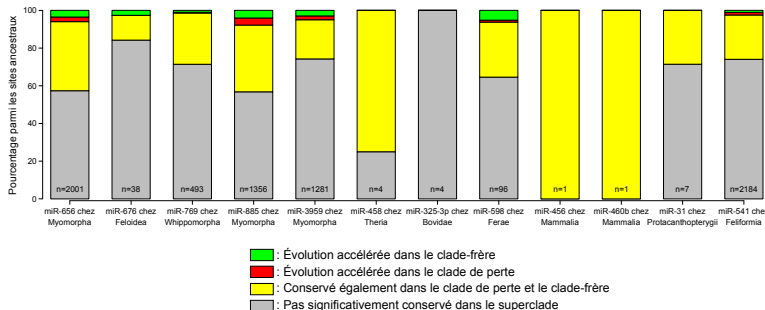
Conclusion

Acknowledgements

Données
supplémentaires



Limitations de la génomique comparative



→ La plupart de ces *seed matches* sont conservés indépendamment du miARN : faux positifs dans l'approche en génomique comparative.

Conclusion

Des propriétés biochimiques uniques (taille, nature des extrémités), qui impliquent des protocoles expérimentaux et informatiques dédiés.

Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Conclusion

Des propriétés biochimiques uniques (taille, nature des extrémités), qui impliquent des protocoles expérimentaux et informatiques dédiés.

Sensibilité du Small RNA-Seq : parfois source de confusion.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Conclusion

Des propriétés biochimiques uniques (taille, nature des extrémités), qui impliquent des protocoles expérimentaux et informatiques dédiés.

Sensibilité du Small RNA-Seq : parfois source de confusion.

Fixation des cibles : reconnaissance *digitale* ($\neq$ facteurs de transcription, *RNA-binding proteins* : analogique), qui facilite la prédiction.

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Conclusion

Des propriétés biochimiques uniques (taille, nature des extrémités), qui impliquent des protocoles expérimentaux et informatiques dédiés.

Sensibilité du Small RNA-Seq : parfois source de confusion.

Fixation des cibles : reconnaissance *digitale* ($\neq$ facteurs de transcription, *RNA-binding proteins* : analogique), qui facilite la prédiction.

Des faux positifs aussi bien dans les approches expérimentales qu'informatiques pour identifier les cibles fonctionnelles.

Acknowledgements



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Acknowledgements



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

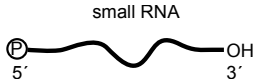
Conclusion

Acknowledgements

Données
supplémentaires



Small RNA-Seq



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

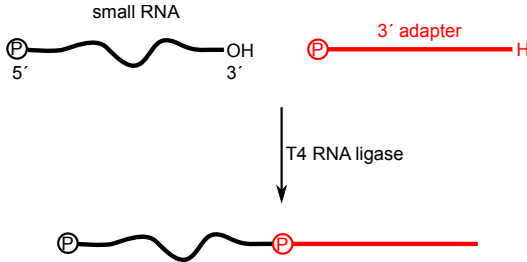
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Small RNA-Seq



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

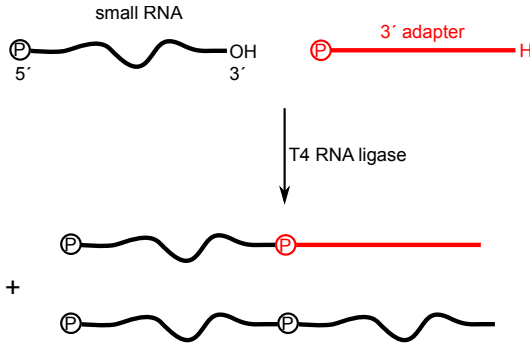
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Small RNA-Seq



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

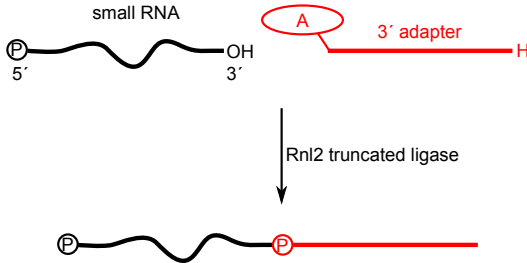
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Small RNA-Seq



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

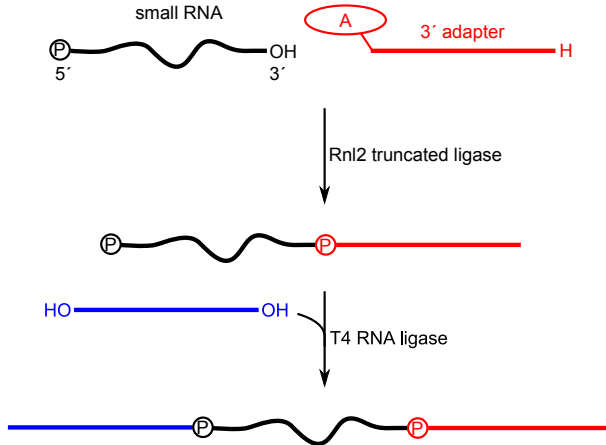
Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires

Small RNA-Seq



Analyses
bio-informatiques
de la biologie des
microARN

Particularités des
microARN

Répression des
cibles

Identification à
haut débit des
cibles de
microARN

Limitations de la
génomique
comparative

Conclusion

Acknowledgements

Données
supplémentaires