

ΠΕΡΙΛΗΨΗ ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ ΙΩΑΝΝΗ ΒΑΤΣΕΛΑ (ΕΝΟΤΗΤΑ IV)

Η ρύθμιση της γονιδιακής έκφρασης είναι μια θεμελιώδη κυτταρική λειτουργία που καθορίζει ζωτικές βιολογικές διεργασίες, όπως η κυτταρική διαφοροποίηση, η ανάπτυξη των πολυκύτταρων οργανισμών και οι κυτταρικές αποκρίσεις σε περιβαλλοντικά ερεθίσματα και παθογόνα. Η ικανότητα διαφορετικών κυττάρων να προγραμματίζουν το, ταυτόσημο κατά τα άλλα, γονιδίωμα τους και να επιλέγουν τα κατάλληλα γονίδια που θα εκφραστούν σε κάθε χρονική στιγμή και μετά από συγκεκριμένα ερεθίσματα, είναι δομικής σημασίας για τους ανώτερους πολυκύτταρους οργανισμούς. Για το σκοπό αυτό, οι ευκαρυωτικοί οργανισμοί έχουν αναπτύξει ποικίλους μηχανισμούς, οι οποίοι βασίζονται σε γενετικούς, επιγενετικούς, και σχετιζόμενους με τη χρωματίνη παράγοντες. Η χρωματίνη είναι μια δυναμική νουκλεοπρωτεϊνική δομή, μέσω της οποίας διαμορφώνεται και διαχειρίζεται το γενετικό υλικό, έτσι ώστε να πραγματοποιούνται διαδικασίες όπως η κυτταρική διαίρεση, η διαφοροποίηση, η επιδιόρθωση του DNA και η ρύθμιση της γονιδιακής έκφρασης. Βασικές επαναλαμβανόμενες μονάδες της χρωματίνης είναι τα νουκλεοσώματα, κάθε ένα από τα οποία αποτελείται από 8 ιστόνες (2 H2A, 2 H2B, 2H3 και 2H4) και DNA μήκους 147 bp τυλιγμένο γύρω από αυτό το οκταμερές των ιστονών. Η παντός είδους τροποποίηση των νουκλεοσωμάτων προσδίδει τα δυναμικά χαρακτηριστικά στη χρωματίνη, τα οποία όντας κληρονομούμενα και με επίδραση στη μεταγραφή των γονιδίων, έχουν σημαντικό ρόλο στον καθορισμό της κυτταρικής ταυτότητας. 3 είναι οι μέχρι στιγμής γνωστοί μηχανισμοί που χρησιμοποιούνται από τα κύτταρα για την αυστηρή ρύθμιση των δυναμικών χαρακτηριστικών των νουκλεοσωμάτων είναι οι μετά-μεταφραστικές τροποποιήσεις των ιστονών, η αναδόμηση της των νουκλεοσωμάτων, και η ανταλλαγή ιστονικών ποικιλομορφών. Οι τελευταίες είναι μη αλληλικές ισομορφές των κανονικών ιστονών, οι οποίες εκφράζονται καθ' όλη τη διάρκεια του κυτταρικού κύκλου, και των οποίων οι ξεχωριστές βιοφυσικές ιδιότητες επιφέρουν σημαντικές τροποποιήσεις στη δομή και τις ιδιότητες της χρωματίνης, μετά από την ενσωμάτωσή τους στα νουκλεοσώματα στη θέση των κανονικών τους ισοδύναμων. Τέτοιες ποικιλομορφές είναι οι H2A.Z, H2A α , H2ABBD και macroH2A (της H2A), και οι H3.3 και CENPA (της H3).

Η macroH2A είναι μια ασυνήθιστη ποικιλομορφή με μέγεθος ~3 φορές μεγαλύτερο της κανονικής H2A, η οποία είναι υψηλά συντηρημένη μεταξύ των σπονδυλωτών. Το αμινοτελικό της άκρο παρουσιάζει ~65% ομοιότητα με την κανονική H2A και είναι ενωμένο με μια μεγάλη μη ιστονική περιοχή, γνωστή και ως macro-δομή. Υπάρχουν 3 διαφορετικά παράγωγα της macroH2A. Οι macroH2A1.1 και macroH2A1.2 προέρχονται από εναλλακτικό μάτισμα του γονιδίου *H2AFY*, ενώ η macroH2A2 κωδικοποιείται από διαφορετικό γονίδιο, και εκφράζεται σε πολύ μικρότερη ποσότητα. Η macroH2A φαίνεται να έχει σημαντικό ρόλο στην καταστολή της γονιδιακής έκφρασης. Έχει δείχθει ότι η δομή των νουκλεοσωμάτων που περιέχουν macroH2A είναι τέτοια που παρεμποδίζει την πρόσδεση μεταγραφικών παραγόντων σε γειτονικές *in cis* ρυθμιστικές περιοχές, ενώ παράλληλα προσδίδουν ανθεκτικότητα στην αναδόμηση τους από σύμπλοκα όπως το SWI/SNF. Επιπλέον, εντοπίζεται κυρίως σε μεταγραφικά ανενεργές χρωματινικές περιοχές όπως το ανενεργό X χρωμόσωμα των θηλυκών θηλαστικών, η περικεντρομερική χρωματίνη, και ετεροχρωματινικές περιοχές που σχετίζονται με τη γήρανση (senescence). Πάντως η αναλογία των macroH2A προς τα κανονικά νουκλεοσώματα είναι περίπου 1 προς 500, και η κατανομή τους φαίνεται να είναι σε γενικές γραμμές ομοιόμορφη κατά μήκος του γονιδιώματος. Πρόσφατα αποδείχτηκε στο εργαστήριο μας ότι η macroH2A1 είναι υπεύθυνη για την ιστό-ειδική καταστολή της επαγόμενης έκφρασης γονιδίου της ιντερλευκίνης 8 (*IL-8*). Σε επιθηλιακά κύτταρα (HeLa), ο υποκινητής/ενισχυτής του γονιδίου δεν καλύπτεται από νουκλεόσωμα, και η μεταγραφή επάγεται απρόσκοπτα, μετά από κατάλληλο ερέθισμα (ϊική μόλυνση). Σε Β-λεμφοκύτταρα (NAMALWA) ωστόσο, ένα macroH2A1 νουκλεόσωμα καλύπτει την αντίστοιχη περιοχή, εμποδίζοντας τη μεταγραφή του γονιδίου.

Στόχος της παρούσης εργασίας είναι η διερεύνηση ενός πιθανού ευρύτερου ρόλου της macroH2A1, στον επιγενετικό καθορισμό ιστό-ειδικών προγραμμάτων γονιδιακής έκφρασης. Για το σκοπό αυτό, με τη χρήση τεχνολογιών παρεμπόδισης του RNA με βάση τους λέντι-ιούς, δημιουργήθηκαν νέες σταθερές κυτταρικές σειρές HeLa και NAMALWA με μειωμένη έκφραση της macroH2A1. Χωρίς ή/και μετά από μόλυνση τους με ιό Sendai, από αυτές απομονώθηκε συνολικό mRNA, το οποίο υβριδοποιήθηκε σε πλακάκια μικροσυστοιχιών έκφρασης της Affymetrix, με στόχο τον προσδιορισμό του συνόλου των γονιδίων, επαγόμενων και μη, η έκφραση των οποίων εξαρτάται από τη macroH2A1, και στις δυο κυτταρικές σειρές.

Αρχικά, κατασκευάστηκαν βασισμένοι σε λέντι-ιούς πλασμιδιακοί φορείς (pLL3.7) που εκφράζουν μόρια shRNA ενάντια στη macroH2A1, και φορείς αναφοράς (control) που εκφράζουν μη λειτουργικά μόρια shRNA (scramble), και κατόπιν συγκροτήθηκαν αντίστοιχα λέντι-ιοσωμάτια. Μετά τη μόλυνση, οι λέντι-ιοί ενσωματώνονται στο γονιδίωμα των ξενιστών, με αποτέλεσμα τη μακροχρόνια και λειτουργική αποσιώπηση του γονιδίου-στόχου, μέσω ενεργοποίησης του ενδογενούς μονοπατιού παρεμπόδισης του RNA. Οι φορείς αυτοί εκφράζουν παράλληλα και την πρωτεΐνη EGFP, επιτρέποντας έτσι τη διάκριση των μολυσμένων κυττάρων με κυτταρομετρία ροής (FACS). Επιπλέον, κατασκευάστηκαν παρόμοιοι φορείς, μετά από σειρά διαδοχικών βημάτων κλωνοποίησης και με τη χρήση χαρακτηριστικών στοιχείων από διάφορους πλασμιδιακούς φορείς (pLL3.7, pIRES2EGFP, pCDNA3.1myc/his), οι οποίοι εκφράζουν σημασμένη με ετικέτες myc και his macroH2A1, παράλληλα με την EGFP, καθώς και οι αντίστοιχοι φορείς αναφοράς που εκφράζουν μόνο EGFP (IRES).

Οι φορείς χρησιμοποιήθηκαν για τη μόλυνση κυττάρων HeLa και NAMALWA, κάτω από τις βέλτιστες συνθήκες, όπως αυτές προσδιορίστηκαν από διαδοχικές αναλυτικές μολύνσεις. Εν συνεχεία, με ανάλυση FACS αποκτήθηκαν κυτταρικοί πληθυσμοί αποτελούμενοι σχεδόν αποκλειστικά από μολυσμένα κύτταρα, στους οποίους και εντοπίζονται κατά συνέπεια πολύ χαμηλά επίπεδα της macroH2A1. Μετά από την επιβεβαίωση της εξειδίκευσης και αποτελεσματικότητας όλων των διαφορετικών λέντι-ιών σε όλους τους διαφορετικούς κυτταρικούς πληθυσμούς, με ανοσο-αποτυπώσεις κατά western και αντιδράσεις RT-PCR, από όλα τα κύτταρα απομονώθηκε RNA πριν και μετά από επαγωγή τους με ιό Sendai για 6 ώρες. RNA απομονώθηκε επίσης και από αγρίου τύπου κύτταρα, πριν, και μετά από 6 και 12 ώρες επαγωγής με ιό Sendai. Τα RNA χρησιμοποιήθηκαν για υβριδοποίηση στα πλέον περιεκτικά πλακάκια μικροσυστοιχιών HG 113 (Human Genome 133) plus 2.0 της Affymetrix. Όλες οι παραπάνω διαδικασίες πραγματοποιήθηκαν από την αρχή εις τριπλούν, ώστε να εξασφαλιστεί η ανάλυση πραγματικών βιολογικών αντιγράφων. Εν συντομία, η ανάλυση των αποτελεσμάτων των μικροσυστοιχιών πραγματοποιήθηκε ως εξής: Κανονικοποίηση και αφαίρεση του θορύβου με τον αλγόριθμο RMA (Robust Multichip Analysis). Η διάκριση των γονιδίων με στατιστικά σημαντικά διαφορετικά επίπεδα έκφρασης μεταξύ των δειγμάτων και των αντίστοιχων κοντρόλ τους (κύτταρα με μειωμένη έκφραση της macroH2A1 με τα scramble, κύτταρα με υπέρ-έκφραση myc/his macroH2A1 με IRES, και αγρίου τύπου

μολυσμένα με Sendai κύτταρα με μη μολυσμένα) πραγματοποιήθηκε με τον αλγόριθμο SAM (Significant Analysis of Microarray), με κριτήρια αποδεκτού σφάλματος (FDR - False Discovery Rate) <5%, και ελάχιστο βαθμό μεταβολής της έκφρασης το 1,5. Ο χαρακτηρισμός και η αντιστοίχιση των γονιδίων με βιολογικές λειτουργίες και κυτταρικά μονοπάτια εκπληρώθηκε με τη χρήση του λογισμικού EASE, ενώ οι πληροφορίες αντλήθηκαν από τις βάσεις δεδομένων της Affymetrix και DAVID.

Η σύγκριση των επαγόμενων από ιό γονιδίων των δυο κυτταρικών σειρών, στις 6 και στις 12 ώρες μετά από τη μόλυνση, αποκάλυψε εκτεταμένες ιστό-ειδικές διαφορές στην μεταγραφική ενεργοποίηση, πρωτίτως στο νωρίτερο χρονικό σημείο (6 ώρες), στο οποίο ξεκινάει η μεταγραφή της IFN β και στους 2 κυτταρικούς τύπους. Στο αργότερο χρονικό σημείο (12 ώρες), οι διαφορές αυτές αμβλύνονται, καθώς η αντί-ϊική δράση της IFN β , σε αντίθεση από ότι φαίνεται με το ιϊκό ερέθισμα, προκαλεί παραπλήσιες γονιδιακές αποκρίσεις στα διαφορετικά κύτταρα, ενώ χαρακτηριστικό είναι το γεγονός ότι το προφίλ γονιδιακής επαγωγής (τόσο αριθμός των γονιδίων, όσο και οι επηρεαζόμενες βιολογικές λειτουργίες) των HeLa τείνει προς αυτό NAMALWA. Τα τελευταία, ως Β-λεμφοκύτταρα, εμφανίζουν παρόμοιο προφίλ γονιδιακής απόκρισης στις δυο χρονικές στιγμές, εστιασμένο μάλιστα στην αντί-ϊική απόκριση. Ένας αριθμός 66 γονιδίων φαίνεται να επάγονται σε όλα τους διαφορετικούς κυτταρικούς τύπους και χρονικά σημεία, προφανώς συγκροτώντας έναν κοινό πυρήνα γονιδιακής απόκρισης στην ιϊκή μόλυνση. Αν και στην πλειοψηφία του ο πυρήνας αυτός αποτελείται από γονίδια με γνωστή αντί-ϊική δράση, περιέχει και αρκετά γονίδια με ποικίλες άλλες λειτουργίες. Πέραν όμως του γονιδιακού αυτού πυρήνα, η πλειοψηφία των προσδιορισμένων γονιδίων εμφανίζουν ιστό-ειδικό πρότυπο μεταγραφικής ενεργοποίησης, παρόλο που οι βασικοί επαγόμενοι από ιό μεταγραφικοί παράγοντες (NF- κ B, IRFs and ATF/Jun) επάγονται και στους δυο κυτταρικούς τύπους. Το γεγονός αυτό ενισχύει ακόμα περισσότερο το ρόλο της δυναμικής δομής της χρωματίνης στον καθορισμό ιστό-ειδικών προγραμμάτων γονιδιακής έκφρασης. Πόσο μεγάλη είναι η συμμετοχή της macroH2A1 στο ρόλο αυτό;

Αρχικά, ανακαλύψαμε ότι σε κάθε κυτταρικό τύπο αυξάνεται η έκφραση ενός σχετικά μικρού αριθμού γονιδίων όταν η ποσότητα της macroH2A1 είναι ελαττωμένη. Συγκεκριμένα, προσδιορίστηκαν 209 και 61 τέτοια γονίδια στα NAMALWA και τα HeLa κύτταρα αντίστοιχα. Πέραν της σημαντικής διαφοράς

στον αριθμό, μόνο 2 γονίδια είναι κοινά στις δυο κυτταρικές σειρές, γεγονός που αποδίδει στη macroH2A1 έναν άκρως ιστό-ειδικό ρόλο στον καθορισμό προγραμμάτων γονιδιακής έκφρασης. Το συμπέρασμα αυτό υποστηρίζουν και οι εκτεταμένες διαφορές των βιολογικών λειτουργιών που επιτελούν τα συγκεκριμένα γονίδια σε κάθε κυτταρικό τύπο.

Η σύγκριση των επαγόμενων από ιό γονίδια, η επαγωγή των οποίων ρυθμίζεται από τη macroH2A1 στις δυο κυτταρικές σειρές, απέδωσε εντοπιστικά αποτελέσματα. Τα γονίδια αυτά είναι 507 και 1662 σε HeLa και NAMALWA αντίστοιχα. Και πάλι είναι προφανείς μεγάλες ιστό-ειδικές διαφορές, καθώς η έκφραση μόνο 99 ιό-επαγόμενων γονιδίων απελευθερώνονται και στις δυο σειρές. Περαιτέρω, η σύγκριση των προσδιορισμένων σε κάθε κυτταρική σειρά γονιδίων πριν και μετά τη μόλυνση με ιό, επίσης επιδεικνύει σημαντικές διαφορές μεταξύ των δυο κυτταρικών τύπων. Σε κάθε σύγκριση, οι διαφορές των βιολογικών λειτουργιών που επηρεάζονται από τα προσδιορισμένα γονίδια, τονίζουν περαιτέρω την ιστό-ειδική δράση της macroH2A1. Φαίνεται λοιπόν πως ενώ η μόλυνση με ιό επάγει τη μεταγραφή 946 και 559 γονιδίων σε HeLa και NAMALWA αντίστοιχα, η απουσία της macroH2A1 απελευθερώνει την επαγωγή 507 και 1662 γονιδίων αντίστοιχα. Επιπλέον, συγκρίνοντας τον αριθμό των ιό-επαγόμενων γονιδίων μεταξύ των κυττάρων αγρίου τύπου και αυτών με ελαττωμένη ποσότητα macroH2A1, προκύπτουν 28 κοινά μεταξύ τους γονίδια στα HeLa και 290 στα NAMALWA, κάτι που υπονοεί ότι τα γονίδια αυτά επάγονται από ιό, σε αγρίου τύπου κύτταρα, μετά από την απομάκρυνση της macroH2A1 από κάποια ρυθμιστική περιοχή τους.

Συνολικά, τα προαναφερθέντα αποτελέσματα δείχνουν ότι η macroH2A1 καταστέλλει την έκφραση λίγων γονιδίων σε αμόλυντα κύτταρα, αλλά έχει σημαντικό ρόλο στην αποτροπή ή/και τη λεπτομερή ρύθμιση της επαγωγικής ενεργοποίησης μεγάλου αριθμού γονιδίων. Και στις δυο περιπτώσεις, η επίδραση της macroH2A1 εμφανίζει ξεκάθαρο ιστό-ειδικό πρότυπο, και μάλιστα με αρκετά μεγαλύτερη βαρύτητα στα πιο διαφοροποιημένα και εξειδικευμένα Β-λεμφοκύτταρα (NAMALWA). Το τελευταίο επιβεβαιώνεται και από το γεγονός ότι στα κύτταρα NAMALWA, σε αντίθεση με τα HeLa, με μειωμένη έκφραση της macroH2A1 επάγεται απόπτωση, με αργή και προοδευτική εξέλιξη, όπως αποδεικνύεται με πειράματα χρώσης αννεξίνης-VE, με ή χωρίς την επίδραση του αποπτωτικού παράγοντα CHX.

Λόγω της εμφάνισης αποίπτωσης στα NAMALWA κύτταρα με μειωμένη ποσότητα macroH2A1, επιλέχθηκαν για περαιτέρω ανάλυση γονίδια που προσδιορίστηκαν στις δυο κυτταρικές σειρές, και τα οποία όπως προκύπτει από τη γονιδιακή τους οντολογία συμμετέχουν σε αποιπωτικά μονοπάτια. Αρχικά, πειράματα RT-PCR επιβεβαίωσαν τα αποτελέσματα των μικροσυστοιχιών DNA, και κατόπιν διερευνήθηκε με πειράματα ανοσοκατακρήμνισης χρωματίνης (ChIP) η φυσική παρουσία macroH2A1 στην ευρύτερη περιοχή των υποκινητών των γονιδίων αυτών. Στην πλειοψηφία των υπό διερεύνηση γονιδίων εντοπίστηκε macroH2A1. Μάλιστα με πειράματα ChIP επαληθεύτηκε και η απομάκρυνση της macroH2A1 από τις περιοχές αυτές, σε όσα γονίδια επάγονται μετά από μόλυνση με ιό στα κύτταρα αγρίου τύπου, σε αντίθεση με αυτά που παραμένουν μεταγραφικά ανενεργά, στα οποία η macroH2A1 παραμένει στη θέση της και μετά την ιική μόλυνση.

Ένα εντυπωσιακό εύρημα είναι ότι, παρά τον πολύ μεγαλύτερο αριθμό γονιδίων που ρυθμίζονται από τη macroH2A1 στα NAMALWA, η συνολική ποσότητα της πρωτεΐνης που αποτίθεται στη χρωματίνη είναι πολύ μικρότερη στα κύτταρα αυτά, από ότι στα HeLa. Ακόμα πιο ενδιαφέρον είναι το εύρημα ότι η υπέρ-έκφραση της *myc/his* macroH2A1, η οποία επίσης αποτίθεται κανονικά στη χρωματίνη, ενώ επηρεάζει δραματικά τη γονιδιακή έκφραση στα HeLa, δεν έχει καμία απολύτως επίδραση στα NAMALWA. Οι παραπάνω παρατηρήσεις, σε συνδυασμό με τον εμφανώς σημαντικότερο αντίκτυπο που έχει η έλλειψη της macroH2A1 στα NAMALWA, υπονοούν την ύπαρξη αυστηρότερων μηχανισμών εναπόθεσης των macroH2A1 νουκλεοσωμάτων στο γονιδίωμα σε πιο διαφοροποιημένους κυτταρικούς τύπους (όπως τα Β-λεμφοκύτταρα), ως μια μέθοδο για την απαιτούμενη λεπτότερη ρύθμιση της γονιδιακής έκφρασης.

Πέραν λοιπόν από τον προσδιορισμό του συνόλου των γονιδίων που ρυθμίζονται από τη macroH2A1 σε διαφορετικές κυτταρικές σειρές, η μελέτη αυτή παρέχει μια σειρά στοιχείων που υποστηρίζουν έναν ιστό-ειδικό ρόλο της macroH2A1 στην επιγενετική ρύθμιση της γονιδιακής έκφρασης, και κατά συνέπεια του καθορισμού των διαφορετικών λειτουργιών των διαφόρων κυτταρικών τύπων των πολυκύτταρων οργανισμών.

SUMMARY

Regulation of gene expression is a fundamental cellular mechanism that underlies critical biological processes, such as cell growth and differentiation, development of multicellular organisms, and responses to environmental signals and infectious microorganisms. The ability of genetically identical cells to program their genome and select the appropriate set of genes for expression at a given time and under a specific stimulus is of crucial importance for multicellular organisms. Eukaryotes have evolved methodical means to regulate the expression of genes. Most of this is attributed to genetic, epigenetic, and chromatin-associated factors. Chromatin is a highly dynamic nucleoprotein complex through which genetic material is structured and maneuvered to elicit cellular processes such as cell division, differentiation, DNA repair and regulation of transcription. In eukaryotes, the core of this structure is composed of the nucleosomes, or repetitive histone octamer units typically enfolded by 147 bp of DNA and consisting of 2 of each canonical histone H2A, H2B, H3 and H4. Modulation of the fundamental nucleosome units contributes to the dynamic structural characteristics of chromatin, which are inheritable and impact on transcription and, therefore, cellular identity. Multiple mechanisms are utilized by cells to tightly regulate chromatin dynamics, including post-translational histone modifications, chromatin remodeling and histone variant incorporation.

Histone variants are non-allelic isomorphs of the canonical histones that exhibit specific expression, localization, and species-distribution patterns. Histone variants are expressed throughout the cell cycle, in contrast to the restricted to the S phase expression of the canonical histones, and their distinct biophysical characteristics alter the chromatin structure and dynamics, upon their incorporation in the nucleosomes as replacements of their canonical counterparts. Such variants are H2A.Z, H2A_x, H2A_{BBD} and macroH2A (H2A variants), and H3.3 and CENPA (H3 variants), and are all related to specific and distinct functions.

MacroH2A is an unusual variant that is restricted, but highly preserved, amongst vertebrates. It is approximately 3 times the size of canonical H2A and has 2

distinct domains – the N-terminus, which is 65% similar to H2A, and a large non histone-like C-terminus, known as macro domain. There are 3 different subtypes of macroH2A1, macroH2A1.1 and macroH2A1.2 that are encoded by the same gene (*H2AFY*) and are generated by alternative mRNA splicing, and macroH2A2 that is encoded by a different gene and expressed in much lower quantity. MacroH2A is considered to be implicated in transcriptional repression, due to its inhibitory structure and genomic distribution. The unusual structure of the macroH2A-containing nucleosomes is such that interferes with transcription factor binding in nearby binding sites, and also renders resistance to remodeling by chromatin remodeling complexes as SWI/SNF. MacroH2A is enriched in transcriptional inactive regions, such as the inactive X (Xi) chromosome in mammalian female cells, pericentric chromatin, and senescence-associated foci. However, macroH2A-containing nucleosomes occur in every ~500 nucleosomes, and seem to demonstrate uniform distribution amongst different genomic regions in the human genome.

Furthermore, recent experiments conducted in our laboratory have shown that macroH2A1 is responsible for the cell-specific repression of the *IL-8* gene transcriptional activation. In expressing epithelial cells (HeLa), the enhancer/promoter region of the gene is nucleosome-free, whereas in non-expressing B-cells (NAMALWA), a macroH2A1-containing nucleosome masks the critical regulatory region, thus preventing the transcription of the gene. MacroH2A1 is selectively recruited to the promoter in B-cells, via direct interactions with ATF-2 bound to the nearby AP-1 site, due to DNA-induced protein allostery triggered by the specific AP-1 site sequence.

Based on the previous findings, it is reasonable to assume that macroH2A1 may have a wider role in the epigenetic determination of cell-specific gene expression programs. The aim of this work is to investigate such a potential global role of macroH2A1, by identifying the global battery of genes whose expression is regulated by this histone variant, in distinct cell-types. To accomplish that, we established novel, stable, cell lines of both HeLa and NAMALWA cells, in which the endogenous expression of macroH2A is knocked-down, by applying lentivirus-based technologies. Total RNA from these cells was extracted and hybridized to Affymetrix whole human genome expression arrays, in order to identify the pool of genes whose expression is regulated by macroH2A.

The study began with the generation of lentivirus-based vectors (pLL3.7) expressing shRNA against macroH2A1 (both macroH2A1.1 and macroH2A1.2 splicing variants) and the corresponding control vector expressing scrambled shRNA. Lentivirus-delivered shRNAs are capable of specific and efficient gene silencing by inducing the RNAi pathway, in a variety of cell types. Upon infection, lentiviruses are stably integrated into the host genome resulting in the long-term stable expression of the shRNA, under the control of the U6 promoter. The vectors also express EGFP under the control of the CMV promoter, thus allowing isolation of the infected cells by Flow Cytometry (FACS). Furthermore, after a series of cloning steps combining the features of several vectors (pLL3.7, pIRES2EGFP, pCDNA3.1myc/his), we constructed functional lentivirus-based vectors expressing both myc and his tagged macroH2A1 and EGFP, under the control of the CMV promoter, as well as the corresponding control vector expressing only EGFP.

Both HeLa and NAMALWA cells were infected with the viruses under optimal conditions, which were determined after series of titrated infections. Sequential FACS analysis allowed the isolation of cell populations, overall comprising of almost 100% infected cells, thus containing very low to undetectable levels of the macroH2A protein. After checking specificity and efficiency of the appropriate effect in every cell population, with both western blot and RT-PCR, HeLa and NAMALWA cells knocked down for macroH2A together with the cells expressing the scrambled shRNA were either mock or Sendai virus (SV) infected for 6 hours followed by RNA isolation and hybridization to the Affymetrix HG 113 (Human Genome 133) plus 2.0 genechip. The same was carried out also with wild type HeLa and NAMALWA cells, with the addition of another time-course of virus infection for 12 hours, as well as with the macroH2A1 over-expressing HeLa and NAMALWA cells, uninfected with SV. In order to have true biological replicates for the microarray analysis, all of the aforementioned procedures were carried out in distinct triplicates.

The analysis of the microarray data was carried out as follows. Normalization and background subtraction was done using the RMA algorithm (Robust Multichip Analysis). SAM analysis (Significant Analysis of Microarray) was applied to distinguish genes whose expression levels are statistically significantly different between sample (shRNA macroH2A expressing cells) and control (scrambled shRNA) chips, either un-induced or Sendai virus induced. FDR (False Discovery

Rate) <5% was selected as a cutoff and genes with fold change greater than 1.5 (up-regulated) or lower than 1/1.5 (down-regulated) were considered to display differential expression. Analysis of genes in relation to biological processes, molecular functions and pathways was performed after the annotation and classification of genes into pathways and networks by using the EASE software.

Comparison of the virus induced genes between the 2 cell lines, at both 6 and 12 hours after infection, revealed extensive cell-specific differences in the transcriptional response, primarily at the earlier time-point (6 hours), which coincides with the activation of IFN β transcription. At the later time point (12 hours), those differences alleviate, while the overall transcriptional profile (number of genes and gene ontology) of the epithelial cells inclines to that of the more immediate responsive to virus infection B-cells, a result indicating that prior to the exert of IFN β signaling in promoting antiviral response, different cell types response variably to the same external signal (virus infection). Apart from the cell-specific differences, a number of 66 genes expressed in every case have been identified, presumably forming a core of antiviral-response genes common in all cell-types. The biggest part of this antiviral core consists of many known antiviral genes, but it also includes several genes with different functions. However, the majority of the identified inducible genes exhibits tissue specific activation pattern, even though the basic known virus inducible transcription factors (NF-kB, IRFs and ATF/Jun), are expressed in both cell lines. The later result suggests that chromatin dynamics is the major mechanism in determining those cell-specific expression patterns. How big could the part of macroH2A1 be in such mechanisms?

Firstly, we discovered that in each cell type a relatively small number of genes are up-regulated when the concentration of macroH2A was reduced by the siRNA method. More specifically 209 and 61 genes were up-regulated in macroH2A knocked down NAMALWA and HeLa cells, respectively. More importantly, only 2 genes were common in the two cell types, a result that strongly suggests that macroH2A plays a role in establishing cell specific gene expression programs. Furthermore, the Gene Ontology (GO) analysis revealed that different cellular functions are mostly affected by the de-repressed genes, between the 2 cell lines. Most of the up-regulated genes in HeLa cells are involved in signal transduction, whereas in NAMALWA cells the majority of the macroH2A-repressed genes are involved in transcription.

Remarkably, a much bigger number of genes are de-repressed in both macroH2A knock-down cell types when the same experiment was repeated using RNA prepared from virus infected cells. In particular, 507 and 1662 genes in HeLa and NAMALWA cells respectively have been identified as genes whose virus inducible expression is regulated by macroH2A1. Again, tissue specificity is obvious, since only a small number of virus-inducible genes (99) are simultaneously de-repressed in both cell types. Furthermore, comparison of the genes up-regulated in each cell line before and after virus infection revealed again an important difference between the two cell types. In each comparison, differences in the cellular functions affected by the de-repressed genes, emanating from Gene Ontology analysis, further highlight the cell-specific role of macroH2A1. Moreover, whilst virus infection activates 946 and 559 genes in WT HeLa and NAMALWA cells respectively, depletion of macroH2A seems to allow the additional de-repression of 507 and 1662 genes in HeLa and NAMALWA, respectively. However, a comparison between the number of virus-induced genes in WT and macroH2A1 knock down cells, reveals 28 genes in HeLa and 290 genes in NAMALWA that are commonly induced in WT and macroH2A1 knock down cells, suggesting that these genes are induced after loss of macroH2A1 from their regulatory region, upon virus infection.

In total, the aforementioned results show that macroH2A1 suffices to repress only a small number of genes in uninfected cells, but has a major role in restricting and fine-tuning the virus-infected gene expression program in both cell types. In each case the effect of macroH2A1 is cell type-specific, and with much higher impact in the more differentiated B-cells (NAMALWA). This is also supported by the observation that knockdown of macroH2A1 in NAMALWA, but not in HeLa cells induces apoptosis in a slowly progressive manner, as shown by Annexin-VE staining assays, with or without treatment with the apoptosis inducer CHX.

Due to the apparent induction of apoptosis in NAMALWA, the apoptosis-related genes that were up-regulated in both cell types were further investigated with RT-PCR experiments that verified the microarray data. Furthermore, Chromatin Immunoprecipitation (ChIP) assays, verified the physical presence of macroH2A1-containing nucleosomes on the promoter of the majority of these genes in HeLa and/or NAMALWA cells. ChIP experiments also corroborated the loss of macroH2A1 from the promoter of these genes that are virus-induced in WT cells, while showing that macroH2A1 persists in the genes that are not induced, thus

preventing their transcriptional activation. Notably all of the above experiments again demonstrate cell-specific pattern, beginning from the higher number of apoptosis-related genes induced in NAMALWA cells.

An impressive finding is that despite the higher despite the considerably higher number of genes regulated by macroH2A1 in NAMALWA cells, the total amount of the protein available and deposited on the chromatin is much less in MAMALWA, than in HeLa. Strikingly, over-expression of the exogenous myc/his tagged macroH2A1, which is proved to also be deposited on chromatin, dramatically affects global gene expression in HeLa, but has no effect at all in NAMALWA. The above findings, together with the far bigger impact that the knock down of macroH2A1 seems to have on NAMALWA, imply a tighter control of macroH2A1-containing nucleosomes deposition along the genome in more differentiated cell lines, such as B-cells, as a mean of fine-tuning of gene expression regulation.

Taken together, the results, apart from providing a pool of genes regulated by macroH2A1, supply a series of evidence supporting a cell-specific role of macroH2A1 as an epigenetic regulator of global gene expression, and thus a part in the complicated mechanisms determining the establishment of distinct cell types in multicellular organisms.