

	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSIÓ	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 1 de 13

APROVACIÓ

REDACTAT PER	REVISAT PER	APROVAT PER:
Nom: Pilar Mancera Càrrec: Gestora Plataforma Genòmica UAT	Nom: Mònica Anglada Càrrec: Responsable Qualitat Nom: Rosa Prieto Càrrec: Responsable UAT	Nom: Rosa Prieto Càrrec: Responsable UAT
Signatures: <i>(es faran en format digital a l'aplicatiu de gestió documental HACQLT)</i>		

 Vall d'Hebron Institut de Recerca	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSIONÓ	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 2 de 13

Aquest document té com a objectiu proporcionar una guia als usuaris del servei de microarrays d'expressió, incloent la informació més rellevant per dur a terme el disseny experimental, així com la descripció del flux de treball d'aquesta tècnica.

1. INTRODUCCIÓ

La UAT disposa de dues plataformes d'arrays d'Affymetrix-ThermoFisher (**GeneChipSystem 3000** per arrays individuals i **Genetitan** per plaques), que ens permeten adaptar-nos a qualsevol tipus de projecte.



Array individual i equip GeneChip System 3000 (forn d'hibridació, estació fluídica i escàner).



Placa d'arrays i equip GeneTitan

La casa comercial ThermoFisher disposa de diferents tipus d'arrays:

- per l'estudi de l'expressió gènica diferencial: **Clariom S** i **Clariom D** (aquests són els formats més actuals, hi ha altres dissenys més antics que també estan disponibles).

	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSION	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 3 de 13

- per l'estudi de perfils de miRNAs i altres ARNs no codificants: **miRNA 4.0** (format cartutx)/**4.1**(format placa).

Atès que aquests estudis habitualment són complexos, recomanem als investigadors contactar amb nosaltres abans de començar el projecte, per dissenyar la millor aproximació experimental (controls, número de mostres, tipus d'array, etc) amb un pressupost a mida.

2. DISSENY EXPERIMENTAL

En primer lloc, en funció del nombre de mostres i transcrits a estudiar s'ha de determinar quina tècnica és la més apropiada:

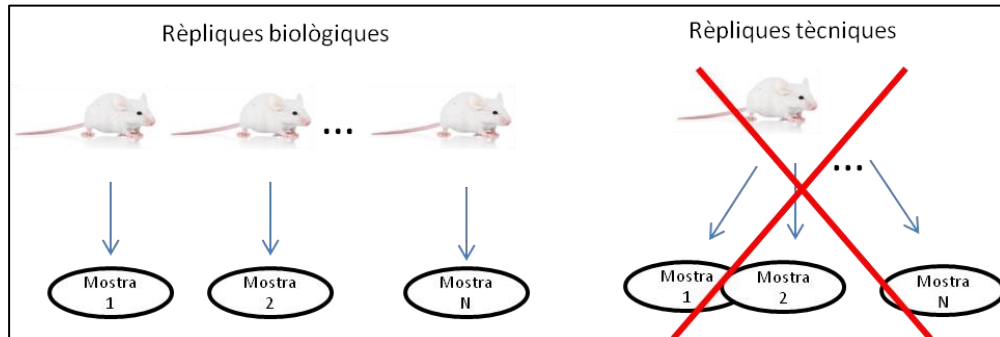
- ✓ Si es tracta d'estudiar l'expressió de **pocs gens en un nombre reduït de mostres**, és aconsellable fer qPCR en un format convencional (plaques de 96 o 384 pouets).
- ✓ Si es vol analitzar un **nombre més elevat de gens o de mostres**, seria adequat fer qPCR utilitzant un sistema de mitjà (TLDA's o arrays de baixa densitat) o d'alt rendiment (per exemple, el sistema Fluidigm o el OpenArray). Altres tecnologies com el Nanostring també podrien ser d'aplicació en aquest cas.
- ✓ Per analitzar el **transcriptoma sencer**, s'haurien de fer servir tecnologies que permeten fer screening, com els microarrays o el RNASeq.

Si es decideix que la tècnica més apropiada són els microarrays, hi ha uns **requisits importants** dels que depèn l'obtenció de resultats fiables i reproduïbles, donada la gran inversió de temps i diners que sovint es fan en aquests experiments:

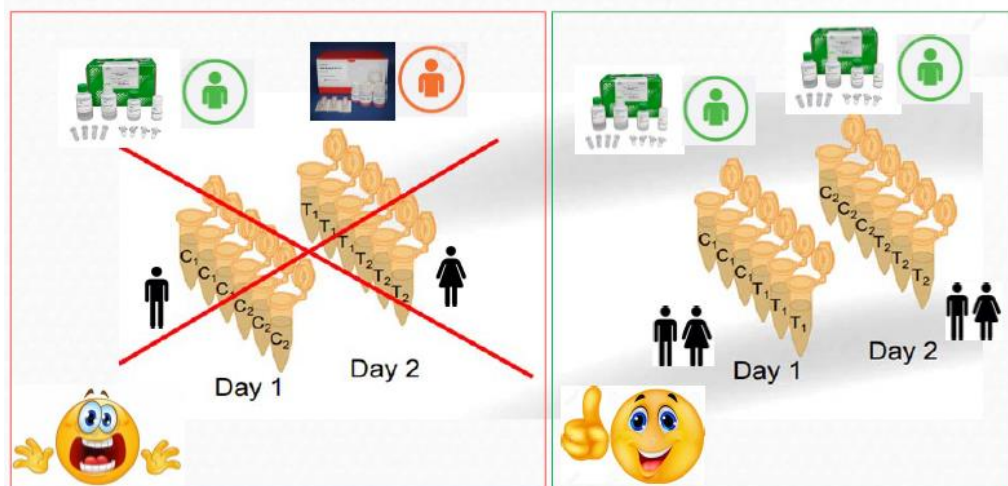
- a) Cal plantejar-se quin és l'**objectiu** de l'experiment, perquè això definirà en gran mesura el disseny de l'experiment (mida de la mostra, tècnica a emplenar, controls, etc). No es podran treure conclusions d'un experiment, si el disseny no és adequat a la hipòtesi de partida.
- b) Una vegada establert l'objectiu, **s'han de tenir en consideració els següents punts**:
 - ✓ És important entendre que la tecnologia d'aquests arrays és molt robusta, així doncs no cal incloure rèpliques tècniques a l'experiment. Però en canvi és fonamental tenir rèpliques biològiques per a cada grup de l'estudi, de manera que siguin una representació el més "real" possible de la seva població o grup d'origen. La variabilitat biològica determinarà la mida mostral: quan més gran sigui la variabilitat esperada, més gran haurà de ser també la representació seleccionada (o, el que és el mateix, la "n"). En línies cel·lulars, és aconsellable una n mínima de 4 mostres per grup. Per animal d'experimentació, hauria de ser més gran, i per humans encara més (perquè la variabilitat individual és més important). Si a priori no es poden estimar les diferències entre dos condicions o grups, una alternativa seria fer un estudi pilot, i a partir d'aquí calcular la

	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSIONÓ	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 4 de 13

n per fer un estudi més gran o validar els resultats mitjançant una altra tècnica (per exemple, qPCR).



- ✓ S'han d'incloure els controls adequats per a cada experiment.
- ✓ No és aconsellable fer “pools” o barreges de l'ARN de diferents mostres per hibridar-lo tot amb un sol array, atès que el resultat obtingut serà degut a la contribució de totes les mostres individuals, i és impossible inferir la contribució de les mostres individuals d'origen. Només es recomana fer “pools” si la quantitat d'ARN obtinguda d'una sola mostra és insuficient per a començar el protocol.
- ✓ És molt important minimitzar tots els factors que introdueixen variacions durant el processament de les mostres, tant durant la recollida, extracció de l'ARN, com del procediment d'amplificació, marcatge i hibridació dels arrays. La recomanació és aleatoritzar tot el que es pugui, i bloquejar tot allò que no es pugui, per evitar efectes “batch” no desitjats que puguin emmascarar la variabilitat biològica real.



- ✓ Pel motiu anterior, no és aconsellable analitzar conjuntament estudis fets de manera independent (només a nivell de llistes de gens diferencialment expressat o si són estudis longitudinals). Tampoc no es poden comparar

	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSIONÓ	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 5 de 13

arrays en format cartutx amb arrays en format placa en un mateix experiment.

- ✓ Per últim, és important definir el mètode més apropiat per a la extracció del ARN, seleccionar el kit d'amplificació en funció de les característiques del material de partida, i el tipus d'array segons les dades que es necessitin obtenir.

Per tal de poder dissenyar un experiment òptim, es recomana fer una reunió amb el personal tècnic de la UAT, a la que també pot participar personal tècnic de la Unitat d'Estadística i Bioinformàtica (UEB).

Una vegada s'han decidit la mida mostral, el disseny experimental, i s'ha seleccionat el tipus d'array i el format (cartutx, placa), la UAT enviarà els pressupostos adients a l'investigador. Si a més de l'execució tècnica de l'experiment es vol sol·licitar el servei d'anàlisi de la UEB, aquesta Unitat enviarà també el seu pressupost.

3. ESTUDIS D'EXPRESSIONÓ GÈNICA DIFERENCIAL:

3.1 Preparació de l'ARN:

El primer pas és preparar l'**ARN total** per fer l'experiment de microarrays.

Per fer l'**extracció de l'ARN**, es recomanen els següents kits:

- QIAGEN: RNeasy Mini Kit, N° Catàleg: 74104.
- Applied Biosystems: MagMAX™ mirVana™ Total RNA Isolation Kit (es necessita un imant específic), N° Catàleg: A27828.
- Invitrogen: PureLink™ RNA Mini Kit, N° Catàleg: 12183018A.
- Invitrogen: TRIzol™ Plus RNA Purification Kit, N° Catàleg: 12183555. Aquest kit incorpora una columna de purificació a la extracció, atès que només el TRIzol és molt eficient però l'ARN extret sol estar contaminat (ratio A260/230 baix).

Tot i que qualsevol kit comercial que ens permeti obtenir un ARN de bona qualitat, en principi, hauria de funcionar bé.

3.2 Avaluació de la qualitat de l'ARN de partida:

Una vegada extret el ARN, cal avaluar la **qualitat** i la **integritat** de les mostres.

Les mostres han de tenir la millor qualitat possible, per tant, ha d'estar lliure de proteïnes, ADN, fenols, etanol i sals. Per avaluar la qualitat utilitzarem els ratios d'absorbància del **NanoDrop®** A260/280 i A260/230. Indiquen un ARN de qualitat acceptable els valor de **A260/280 entre 1.7-2.1**, i de **A260/230 < 2**.

	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSION	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 6 de 13

La integritat de l'ARN és important, ja que la transcripció inversa d' ARNm inicialment degradat pot generar ADNc que li manquen parts de les regions de codificació. Per a la seva avaluació utilitzarem el **Bioanalyzer 2100®** (Agilent) (RNA 6000 Nano Kit o RNA 6000 Pico kit, segons la concentració esperada). El RIN (*RNA Integrity Number*) permet valorar objectivament la integritat de la mostra, en una escala del 1 (totalment degradat) al 10 (ARN intacte). Idealment, s'hauria de treballar amb mostres amb **RIN ≥ 7**, però si això no es possible (per exemple, mostres parafinades), el més important és que totes les mostres tinguin un valor de RIN similar. Actualment existeixen kits per poder treballar amb mostres de RIN baix, com s'indica a continuació.

La concentració inicial per començar el protocol es calcularà a partir de les mesures obtingudes pel bioanalyzer.

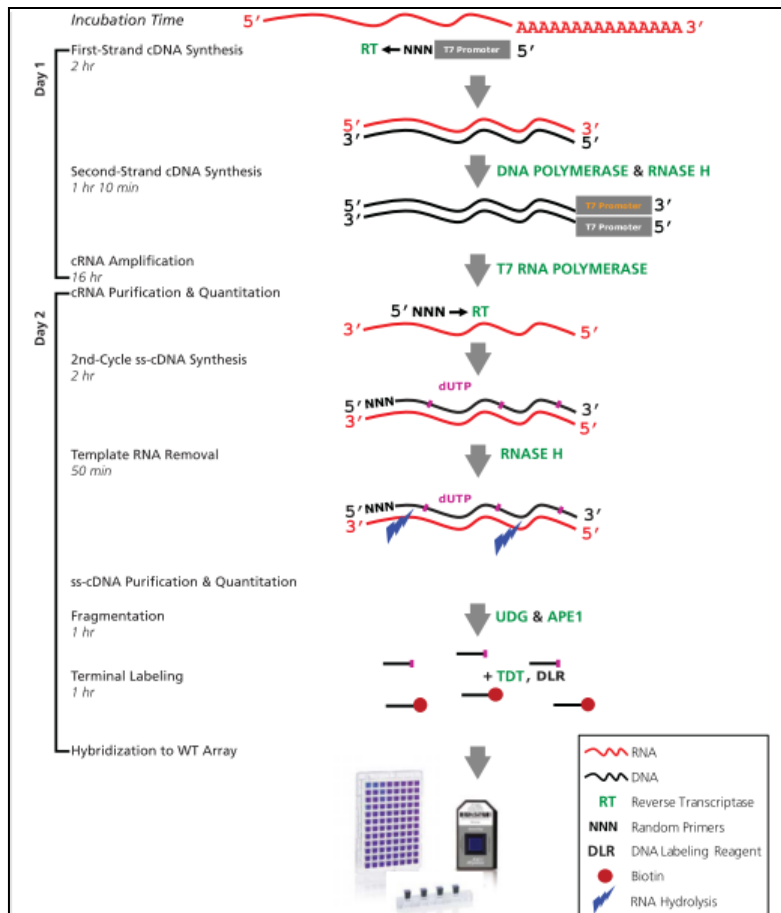
3.3 Tipus de kits d'amplificació segons la quantitat i tipus de mostra:

GeneChip® WT Plus Reagent Kit

Està optimitzat per treballar amb una ampla varietat de mostres, on s'inclouen: teixit fresc o congelat, cèl·lules, línies cel·lulars i sang total. Es pot començar amb **50ng** d'ARN total fins a un màxim de 500ng.

RNA Input	Total RNA
Recommended	100 ng
Minimum	50 ng
Maximum	500 ng

El flux de treball del kit és:



GeneChip® WT Pico (per arrays Clariom D) o Pico Reagent Kit (per arrays Clariom S)

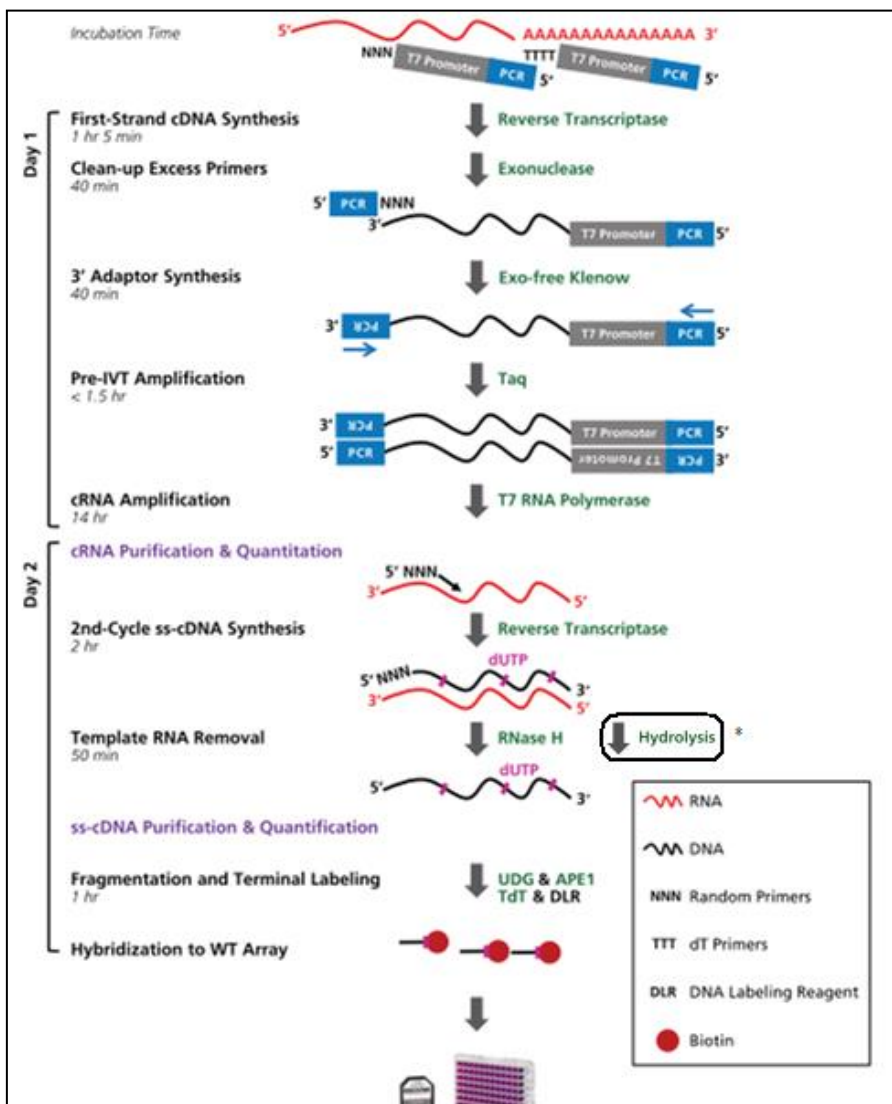
Aquests kits estan especialment dirigits a mostres d'ARN parcialment degradat o mostres **molt compromeses** (RIN <4), per exemple **mostres parafinades (FFPE)**. La diferència més important respecte al kit *WT Plus* és que l'amplificació de l'ARN s'aconsegueix fent una PCR (de un número baix de cicles) i a continuació una amplificació lineal fent una transcripció in vitro (IVT). Per aquest motiu també permet començar a partir d'una quantitat d'ARN total inferior al kit *WT Plus*. És a dir, aquests kit són adients per a mostres d'ARN amb baixa concentració o baixa qualitat.

Estan optimitzats per treballar amb una gran varietat de mostres, on s'inclouen: teixit fresc o congelat, cèl·lules, línies cel·lulars, sang total i **mostres parafinades (FFPE)**.

Es pot començar amb **100pg** d'ARN fins a un màxim de 10ng, si son mostres de teixits o cèl·lules. I, si són mostres parafinades s'haurà de començar amb **500pg** fins a un màxim de 50ng.

RNA Input	Total RNA from Fresh-Frozen Cells or Tissues	Total RNA from Formalin-Fixed, Paraffin-Embedded Tissues
Minimum	100 pg	500 pg
Recommended	500 pg – 10 ng	2 ng – 50 ng
Maximum	10 ng	50 ng

El flux de treball d'aquest kit és:



*Diferència entre WT Pico i Pico està en la forma d'eliminar el "template ARN", el **Pico** en lloc d'utilitzar la RNase H fa una hidròlisi.

En resum, segons el tipus de mostra i la quantitat d'ARN total de partida s'utilitzarà un kit o un altre:

Sample type	Fresh/frozen tissue		FFPE
	Whole blood		
Input	50-500 ng	0.1-10 ng	0.5-50 ng
Clariom™ D arrays	GeneChip® WT PLUS Reagent Kit	GeneChip® WT Pico Kit	
Clariom™ S arrays	GeneChip® WT PLUS Reagent Kit	GeneChip® Pico Kit	

3.4 Tipus d'arrays d'expressió gènica diferencial:

Les versions d'arrays més actuals són **Clariom S** i **Clariom D**, per humà, ratolí i rata. A més hi ha versions d'arrays amb anotacions més antigues (Gene Array, HTA...), i arrays per altres espècies model. Per projectes nous recomanem fer servir els Clariom, ja que l'anotació està feta amb bases de dades més recents i el disseny s'ha millorat en comparació amb versions anteriors (a més, són més econòmics). Les versions antigues les recomanem només en projectes de continuïtat.

- **Clariom S**, que és una versió reduïda del Clariom D, conté sondes només pels exons constitutius dels més de 20.000 gens anotats, i per tant és adequat per estudis senzills d'expressió diferencial:

Content summary	Human	Mouse	Rat
Genes*	>20,800	>22,100	>22,900
Transcripts*	>337,100	>150,300	>129,800
Total probes*	>211,300	>221,900	>231,800
Probes targeting genes*	>205,800	>221,300	>229,500
Probe length (bases)	25	25	25
Probe feature size	5 µm	5 µm	5 µm
Background probes	Antigenomic set	Antigenomic set	Antigenomic set
Probe orientation**	Antisense	Antisense	Antisense

- **Clariom D**, conté una elevada densitat de sondes per a tot el transcriptoma, incloent la detecció de lncRNA i de formes de "splicing" alternatiu. També permet detectar transcrits de baixa expressió.

Content summary	Human	Mouse	Rat
Genes*	>134,700	>66,100	>68,900
Transcripts*	>542,500	>214,900	>495,200
Exons*	>948,300	>498,500	>320,400
Exon-exon splice junctions*	>484,900	>282,500	>293,700
Total probes*	>6,765,500	>6,022,300	>5,946,400
Probes targeting exons*	>4,781,200	>4,895,600	>4,780,700
Probes targeting exon-exon splice junctions*	>1,984,300	>1,126,700	>1,165,700
Probe length (bases)	25	25	25
Probe feature size	5 µm	5 µm	5 µm
Background probes	Antigenomic set	Antigenomic set	Antigenomic set

4. ESTUDI DE PERFILS DE miRNAS I ALTRES ARNs NO CODIFICANTS

4.1 Preparació de l'ARN:

Per aïllar l'ARN es farà servir qualsevol kit d'extracció d'**ARN total**, però que **retingui** també els **ARN de baix pes molecular**. Alguns kits comercials que son compatibles amb el FlashTag Biotin HSR són:

- Applied Biosystems: mirVana™ miRNA Isolation Kit
- Applied Biosystems: RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE
- QIAGEN: miRNeasy Mini Kit
- Invitrogen: PureLink™ miRNA Isolation Kit
- Invitrogen: TRIzol® reagent (total RNA only) with additional overnight -20°C precipitation step during isopropanol precipitation¹

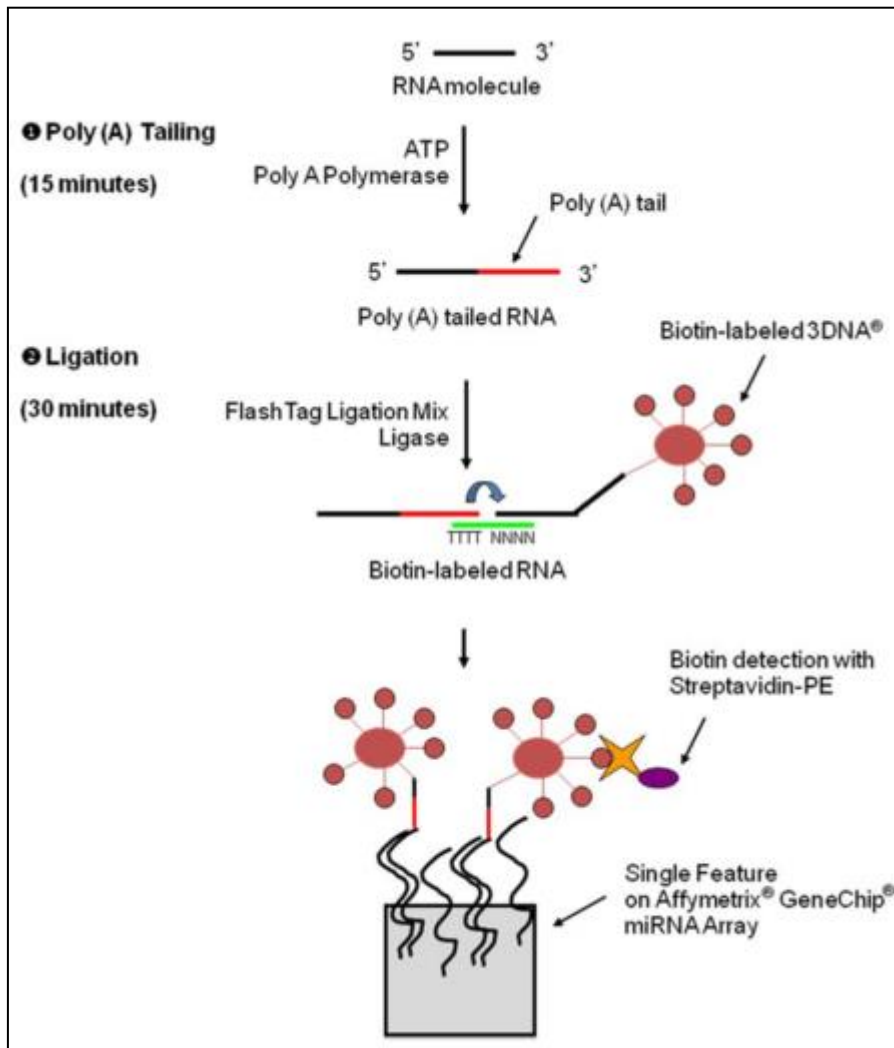
La quantitat mínima d'**ARN total** que es necessita es de **130ng**. Es recomana fer la quantificació de l'ARN mitjançant el Quant-iT® Ribogreen RNA Assay kit (Invitrogen) o NanoDrop® ND-1000.

RNA Sample	Input for FlashTag Biotin HSR Labeling for miRNA 400/169 Format Arrays (miRNA 1.0 and 2.0 Arrays)	Input for FlashTag Biotin HSR Labeling for miRNA 100 Format Arrays (miRNA 3.0 and later designs)
Total RNA containing LMW RNAT	100 - 1000 ng total RNA	130 - 1000 ng total RNA
Enriched LMW RNA, quantitated	100 - 400 ng LMW RNA	130 - 400 ng LMW RNA
Enriched LMW RNA, not quantitated	Enriched from 100 - 1000 ng total RNA	Enriched from 130 - 1000 ng total RNA

	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSIONÓ	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 11 de 13

4.2 Kit de marcatge:

Per l'estudi dels miRNAs es farà servir un kit de marcatge, el **FlashTag Biotin HSR RNA Labelling kit**. Marcarà qualsevol mostra d'ARN, incloent ARN total, ARN severament degradat, ARN vegetal o ARN de baix pes molecular. El flux de treball per aquest kit és el següent:



Per tal d'assegurar-nos que el marcatge ha funcionat correctament, es fa un control de qualitat, el **ELOSA** (Enzyme Linked Oligosorbent Assay). És un assaig colorimètric, on un control positiu i un negatiu, confirmarà que l'assaig ha funcionat (canviant de color, si hi ha ARN marcat serà blau i si no es quedarà blanc).



Control Negatiu

Control Positiu

4.3 Tipus d'array:

Els arrays que s'utilitzen són els miRNA 4.0 si es format cartux o miRNA 4.1 si es format placa. En un mateix array s'analitzen totes les espècies (humà, ratolí i rata, o qualsevol miRNA).

Array content description	
miRBase	Release 20
Ensembl (BioMart Export)	Release 73
snoRNAbase	Version 3
mirTarbase	Release 4.5
MicroCosm Targets	11/2013
Organisms (including viruses)	203
Total mature miRNA probe sets	30,434
Probes/probe set for mature miRNA	9
Human mature miRNA probe sets	2,578
Mouse mature miRNA probe sets	1,908
Rat mature miRNA probe sets	728
Human snoRNA and scaRNA probe sets	1,996
Human pre-miRNA probe sets	2,025
Mouse pre-miRNA probe sets	1,255
Rat pre-miRNA probe sets	490

5. ANÀLISI DE LES DADES OBTINGUES

Una vegada realitzada la tècnica, la UAT durà a terme un primer control de qualitat per revisar que el protocol experimental s'ha desenvolupat correctament. Si tot és correcte, l'investigador rebrà un enllaç per descarregar els arxius .cel generats.

Cada investigador pot analitzar les seves dades fent servir algun software específic (per exemple el TAC, Transcriptome Analysis Console Software, que es pot descarregar de la web de Thermo Fisher). Alternativament, si s'ha demanat l'anàlisi a la Unitat d'Estadística i Bioinformàtica, la UAT els transferirà els resultats directament.

 Vall d'Hebron Institut de Recerca	GUIA PER A USUARIS DE MICROARRAYS D'EXPRESSIÓ	Codi: VHIR-UAT-DOC-022	Revisió: 01
		Data de redacció: 30/04/2021	Pàgina: 13 de 13

6. VALIDACIÓ DELS RESULTATS

La tècnica dels microarrays és molt robusta, però degut a que és una tècnica d'alt rendiment, el propi disseny dels arrays implica una optimització general de les sondes i això, conjuntament amb el fet de generar milers de mesures per cada mostra, pot generar falsos positius i negatius en fer l'anàlisi dels gens diferencialment expressats. Per aquest motiu, es molt recomanable validar una selecció dels transcrits diferencialment expressats per alguna tècnica alternativa, generalment per qPCR. Aquesta tècnica també es pot fer servir per validar l'expressió diferencial dels gens seleccionats en un nombre més gran de mostres (existeixen formats com els arrays de baixa densitat tipus TLDA o els sistemes microfluidics tipus Openarray, Fluidigm, que poden ser molt adients per aquest tipus d'aproximació).