



**HAL**  
open science

# Etude des populations immunitaires régulatrices et inflammatoires dans les cancers colorectaux

Marie Tosolini

► **To cite this version:**

Marie Tosolini. Etude des populations immunitaires régulatrices et inflammatoires dans les cancers colorectaux. Immunologie. Université Pierre et Marie Curie - Paris VI, 2010. Français. NNT : 2010PA066098 . tel-00815336

**HAL Id: tel-00815336**

**<https://theses.hal.science/tel-00815336>**

Submitted on 18 Apr 2013

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

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

**Université Pierre et Marie Curie**

Ecole Doctorale Physiologie et Physiopathologie

**Thèse de doctorat**

Spécialité immunologie

Présenté par

**Marie TOSOLINI**

Pour obtenir le grade de Docteur de l'Université Pierre et Marie Curie

**ETUDE DES POPULATIONS IMMUNITAIRES REGULATRICES ET  
INFLAMMATOIRES DANS LES CANCERS COLORECTAUX**

Soutenue le 25 janvier 2010

Devant le jury composé de:

Dr Jérôme GALON  
Pr Pierre-André CAZENAVE:  
Dr Anne-Marie SCHMITT-VERHULST  
Dr Bernard MAILLÈRE  
Dr Fathia MAMI-CHOUAIB  
Dr Jessica ZUCMAN-ROSSI

Directeur de Thèse  
Président de Jury  
Rapporteur  
Rapporteur  
Examineur  
Examineur

Thèse préparée au laboratoire :

UMRS 872 - Équipe 15  
Immunologie et Cancérologie Intégrative  
Centre de Recherches des Cordeliers  
15, Rue de l'École de Médecine  
75006 Paris

Dirigé par le Docteur Jérôme Galon

# **Remerciements**

*Tout d'abord merci Jérôme de m'avoir accueillie au sein de ton équipe si atypique. J'ai réellement apprécié ce micro-environnement ! Tu m'as offert la possibilité de connaître de nouveaux horizons jusque là insoupçonnés, de grandir et de mûrir. Je reste admirative devant tes idées prolifiques et si justes.*

*Anne, compagne des premiers moments, parfois difficiles. Tu as su me remonter le moral et m'as montré que jamais rien n'était perdu.*

*Mélanie, j'ai trouvé en toi une oreille attentive, une aide sans faille, des minutes culturelles -au sens large- toujours surprenantes. Tu as pimenté ces années de thèse, semant ici et là des moments inoubliables.*

*Stéphanie, ton aide m'a été infiniment précieuse. Ton aide, au travail, bien sûr, mais aussi tous les à-côtés qui font que j'ai toujours le sourire à la fin d'une journée. Y compris nos horaires de déjeuner toujours synchrones ! Quel bonheur d'avoir partagé cette dernière année avec toi !*

*Tessa, à toi Grain de Fantaisie du laboratoire, je ne dirais qu'un seul mot qui résumera tout : "Coin" ! Bernhard et Gabriela vous m'avez fait découvrir une foule de choses dont un univers parallèle, peuplé d'ARACNE, STRING et autres créatures aux noms étranges. Grâce à vous, j'ai moins peur des statistiques !*

*A mes compagnons de route, Julien, Luc, Riad et Amos, merci d'avoir rendu ce laboratoire si vivant. Julien, je ne pourrais jamais oublier tes sifflements précédant toujours ton arrivée dans la pièce ! Je pense que je resterai à jamais traumatisée par tes blagues... Luc, j'ai apprécié tes digressions culturelles toujours étonnantes. Riad, tes points de vue critiques m'ont enrichie. Amos, j'ai apprécié nos débats sur tout et rien à la fois. Je dois également te remercier pour tes précieux conseils pour domestiquer les ordinateurs ! Bien que parfois, ils restent toujours un peu rebelles...*

*Je n'oublie pas non plus tout ceux qui sont passés dans le laboratoire, tels des étoiles filantes. Ils ont tous participé à l'âme de ce laboratoire. Merci à tous ceux qui m'ont soutenu et qui m'ont donné le courage de retourner sur Paris. Vous ne pouvez imaginer à quel point j'ai apprécié ces petites gourmandises venues du sud !*

*Jean-Luc et Jean-Louis merci de m'avoir offert ces moments de détente*

*Et enfin, merci à toute ma famille. Votre soutien tout au long de ma thèse a été une bouée de sauvetage à laquelle je me suis accrochée avec force. Nos discussions et aussi nos longs silences ont été mon étoile du berger. Sans vous, rien n'aurait été possible.*

# Résumé

L'existence d'une immunité anti-tumorale a longtemps été controversée. Aujourd'hui, de nombreuses études ont montré des réponses anti-tumorales spécifiques. Les lymphocytes infiltrant la tumeur (TIL) jouent un rôle dans l'inhibition de la croissance des tumeurs et sont associés à un meilleur pronostic. Cependant, la présence d'une tumeur malgré le système immunitaire met en évidence des mécanismes d'échappement.

J'ai tout d'abord participé à la caractérisation de la valeur pronostic de marqueurs lymphocytaires. La densité de l'infiltrat lymphocytaire T, en particulier les lymphocytes T mémoires, et leurs localisations au niveau du centre et du front de progression de la tumeur sont associées au devenir clinique des patients. Ces paramètres permettent une classification des patients à risque élevé de récurrence bien plus significative que les paramètres anatomopathologiques. En étudiant plus précisément ces paramètres, nous avons montré l'altération progressive des réponses immunitaires anti-tumorales lors de l'envahissement des ganglions ou d'organes à distance.

Je me suis donc intéressée aux marques de cytotoxicité en relation avec les différentes sous-populations T helper. Nous avons à nouveau trouvé des clusters fonctionnels mais seuls les réponses Th1, cytotoxique et Th17 sont associées au pronostic des patients. Les patients avec une forte réponse Th1 ou cytotoxique sont de bon pronostic tandis que ceux avec une forte réponse Th17 sont de mauvais pronostic. La combinaison de ces paramètres est plus discriminante pour déterminer les patients à risque de récurrence.

Les microARNs interviennent dans de nombreux processus biologiques physiologiques (dont le système immunitaire) et physiopathologiques (dont le cancer). J'ai étudié leurs profils d'expression dans des colons sains et tumoraux. Ces profils sont différents, avec des sur-expressions et sous-expressions de certains microARNs. J'ai également observé l'association de certains microARNs au devenir clinique de patients mais aussi à des marqueurs du système immunitaire.

L'ensemble des travaux réalisés montre l'importance de la réaction immunitaire adaptative pour l'établissement du pronostic des patients atteints de cancers colorectaux. Son évaluation en routine permettrait de mieux orienter les patients vers une stratégie thérapeutique adaptée.

# **Abstract**



Anti-tumoral immunity has been argued for ages. Nowadays, a lot of studies demonstrate specific anti-tumoral response. Tumor-infiltrating T cells (TILs) are associated with the inhibition of tumor development and with the clinical outcome. However the presence of a tumor despite the immune system highlights escape mechanisms.

First, I contributed to the characterization of the prognostic value of lymphocyte markers. The density of T lymphocyte and memory T lymphocyte, and their localization in the center and the invasive margin of the tumor, were associated with improved clinical outcome. This immune classification was more significant than the usual parameter to identify patients who relapse. We also showed the impairment of the immune response coordination when lymph nodes or distant organs were invaded. I analyzed the cytotoxic CD8 T cells markers in relation with helper T subpopulations. We found functional clusters that are associated with clinical outcome. Patients with high expression of the Th1/cytotoxic cluster had prolonged disease-free survival. In contrast, patients with high expression of the Th17 cluster had poor prognosis. The combined analysis of cytotoxic and Th17 cluster gave a better discrimination for relapse.

microRNA are involved in a lot of biological process including the immune system or cancer. I studied their profile of expression in normal and tumoral tissues. Their profiles were different with higher and lower level of some microRNAs. I also observed the association of some microRNAs with a particular clinical outcome and with immune system marker.

All this work shows the importance of the immune system in the prediction of clinical outcome. The immune evaluation within the standard procedure would better determine the best therapy.

# Abréviation

ADAR	adenosine deaminases acting on RNA
AGO	Argonaute
AIB1	amplified in breast cancer 1
APC	adenomatous polyposis coli
ARE	AU-rich element
ARNm	ARN messenger
CEA	carcinoembryonnaire antigene
CDKN1A	cyclin dependant kinase inhibitor
CMH	complexe majeur d'histocompatibilité
CMP	common myeloid progenitor
COX2	cyclooxygénase-2
CPA	cellules présentatrices d'antigènes
CSDC2	cold-shock domain-containing protein C2
CT	centre de la tumeur
CTGF	connective tissue growth factor
DGCR8	DiGeorge syndrome Critical Region gene 8
EAE	encéphalomyelite autoimmune expérimentale
EGF	epidermal growth factor
FAP	familial adenomatous polyposis
FLIP	FLICE-like inhibitory protein
FP	Front de progression
FXR1	fragile-X-mental-retardation-related protein 1
GMP	granulocyte-macrophage progenitor
H3K27me3	tri-methyl histone H3 lysine 27
HMGA2	high mobility group A2
HNPCC	hereditary nonpolyposis colorectal cancer
HSP	heat shock proteins
Hur	AU-rich element-binding protein
IDO	indoléamine-2,3 dioxygénase
IFN $\gamma$	interféron- $\gamma$
IgE	Immunoglobuline de classe E
IGF2R	Insuline-like Growth Factor 2 Receptor
IL	interleukine
IM	Invasive Margin
IRAK1	IL-1 receptor-associated kinase I
IRS-1	induced T regulatory
KIRs	Kiler-cell Immunoglobulin-like Receptors
LT	lymphocytes T
M-CSF	macrophage colony-stimulating factor
miRNA	microARN
MSI	microsatelite instability
NFI-A	nuclear factor I-A
NK	natural killer

nTreg	natural T regulatory
PcG	polycomb group
pDC	cellules dendritiques plasmacitoides
PDCD4	programmed cell death gene 4
PI9	serin-Protease Inhibitor 9
PTGS2	prostaglandine E2
RISC	RNA-induced silencing complex
TFH	T follicular helper
Th1	T helper 1
Th17	T helper 2
TIL	lymphocytes infiltrant la tumeur
TLR	Toll like receptor
TMA	Tissue-MicroArray
TNF	tumor necrosis factor
TNF $\alpha$	tumor necrosis factor- $\alpha$
TRAF6	TNF receptor-associated factor 6
TRAIL	TNF-related apoptosis-inducing ligand
Treg	T régulateur
Tsp1	thrombospondin-1

# Table des figures

1	Classification histologique internationale du cancer du colon . . . . .	4
2	Les différentes voies de différenciation des T helper . . . . .	11
3	Les conversions observées des T helper . . . . .	14
4	Les phases de l'immunoediting . . . . .	15
5	Les mécanismes d'échappement de la tumeur au système immunitaire . . . . .	17
6	La transcription et la maturation des microARNs . . . . .	23
7	Mécanismes d'action des microARNs . . . . .	24
8	Variation d'expression des microARNs dans les cellules du système immunitaire inné	27
9	Variation d'expression des microARNs dans les cellules du système immunitaire adaptatif . . . . .	28
10	Effets possibles des microARNs sur les oncogènes et gènes suppresseurs de tumeur .	30
11	Exemple de marquage FoxP3 et son analyse . . . . .	140
12	Courbes de survie sans récurrence associées à la densité de cellules FoxP3+ . . . . .	141
13	Matrice de corrélation des densités des cellules immunitaires . . . . .	142
14	Relation entre la densité de cellules CD3+ et FoxP3+ . . . . .	143
15	Répartition du rapport des densités de cellules FoxP3+ et CD3+ . . . . .	144
16	Courbes de survie sans récurrence associées au ratio des densités de cellules FoxP3+ sur celles des cellules CD3+ . . . . .	145
17	Répartition du rapport des densités de cellules FoxP3+ et IL17+ . . . . .	146
18	Courbes de survie sans récurrence associées au ratio des densités de cellules FoxP3+ sur celles des cellules IL17+ . . . . .	147
19	Deux exemples de microARNs différenciellement exprimés . . . . .	148
20	Comparaison des niveaux d'expression des microARNs dans les stades T . . . . .	150
21	Comparaison des niveaux d'expression des microARNs dans les stades N . . . . .	151
22	Comparaison des niveaux d'expression des microARNs dans les stades M . . . . .	153
23	Courbes de survie sans récurrence de microARNs . . . . .	155

24	Extrait de la matrice de corrélation entre les données sur les microARNs et les cellules immunitaires innées . . . . .	156
25	Extrait de la matrice de corrélation entre les données sur les microARNs et les cellules immunitaires adaptatives . . . . .	156
26	Modèle de l'interaction lymphocyte T/tumeur . . . . .	161
27	Transfection de cellules pour une sur-expression de miR-657 . . . . .	164
28	Modèle d'influence du système immunitaire sur la récurrence . . . . .	168

# Liste des tableaux

2	Exemples d'oncomiRs et leurs cibles . . . . .	31
3	Exemples de microARNs suppresseurs de tumeur et leurs cibles . . . . .	33
4	microARNs significativement exprimés plus fortement dans les tissus tumoraux . . .	149
5	microARNs significativement exprimés plus faiblement dans les tissus tumoraux . .	150
6	microARNs significativement différentiellement exprimés entre les différents stades T ( $p \leq 0,05$ ) . . . . .	151
7	microARNs significativement différentiellement exprimés entre les différents stades N. ( $p \leq 0,05$ ) . . . . .	152
8	microARNs significativement différentiellement exprimés entre les différents stades M. ( $p \leq 0,05$ ) . . . . .	153
9	Cibles potentielles du miR-657 et processus cellulaire associé . . . . .	165

# Table des matières

<b>Introduction</b>	<b>2</b>
<b>Les cancers colorectaux</b>	<b>2</b>
Épidémiologie . . . . .	2
Facteurs de risque . . . . .	2
Classification . . . . .	3
Extension pariétale de la tumeur : le stade T . . . . .	3
Invasion ganglionnaire : le stade N . . . . .	3
Métastases à distance : le stade M . . . . .	5
Facteurs pronostiques . . . . .	5
Les facteurs intrinsèques à la tumeur . . . . .	5
Le micro-environnement de la tumeur . . . . .	6
<b>Immunité anti-tumorale</b>	<b>8</b>
Lymphocytes T et cancer . . . . .	8
Quelques généralités . . . . .	8
Cytotoxicité . . . . .	8
Antigènes tumoraux et activation des lymphocytes . . . . .	9
Les voies de différenciation des lymphocytes T . . . . .	10
Les Lymphocytes T helper 1 . . . . .	10
Les Lymphocytes T helper 2 . . . . .	12
Les Lymphocytes T régulateurs . . . . .	12
Les Lymphocytes T helper 17 . . . . .	12
Les lymphocytes T Follicular Helper . . . . .	13
La plasticité des lymphocytes T Helper . . . . .	13
Modèle d'équilibre/échappement de la tumeur au système immunitaire . . . . .	15
La surveillance du système immunitaire inné . . . . .	15

La surveillance du système immunitaire adaptatif . . . . .	16
Mécanismes d'échappement de la tumeur au système immunitaire . . . . .	16
Complexité du système tumeur/immunité . . . . .	20
<b>Les microARNs</b>	<b>21</b>
Généralités . . . . .	21
Définition . . . . .	21
Historique . . . . .	21
Conservation dans l'évolution . . . . .	22
Les microARNs dans le génome . . . . .	22
Mécanismes de fonctionnement des microARNs . . . . .	22
La transcription et la maturation des microARNs . . . . .	22
L'action des microARNs sur les ARN messagers . . . . .	23
La régulation de l'expression des microARNs . . . . .	25
Les microARNs et le système immunitaire . . . . .	26
Dans le système immunitaire inné . . . . .	26
Dans le système immunitaire adaptatif . . . . .	28
MicroARN et cancer . . . . .	29
Les oncomiRs . . . . .	30
Les microARNs supprimeurs de tumeurs . . . . .	32
<b>Objectifs</b>	<b>36</b>
<b>Résultats</b>	<b>38</b>
<b>Article 1</b>	<b>38</b>
<b>Article 2</b>	<b>45</b>
<b>Article 3</b>	<b>57</b>
<b>Article 4</b>	<b>118</b>
<b>Données non publiées</b>	<b>139</b>
Les lymphocytes T régulateurs . . . . .	139
Matériel et méthodes . . . . .	139



Résultats . . . . .	140
Les microARNs . . . . .	148
Matériel et Méthodes . . . . .	148
Résultats . . . . .	148
<b>Discussion</b>	<b>158</b>
La réaction immunitaire dans les cancers colorectaux	158
Les microARNs	163
<b>Conclusion</b>	<b>167</b>
<b>Annexes</b>	<b>179</b>
Annexe 1	179
Annexe 2	183

# **Introduction**

# Les cancers colorectaux

## Épidémiologie

Les cancers colorectaux sont au premier rang des cancers tous sexes confondus. 945 000 nouveaux cas par an sont détectés dans le monde, soit 9,4% des cancers totaux. Ils sont la cause de 492 000 décès par an soit 7,9% des décès mondiaux dus au cancer [1]. Ce type de cancer est plus commun dans les pays développés que dans ceux en voie de développement. Cependant le taux de survie à 5 ans est plus important dans les pays développés (60% au Etats Unis contre moins de 40% dans des pays en développement).

Chaque année en France, près de 36 000 nouveaux cas et 15 000 décès associés au cancer du colon sont détectés. Son incidence dans la population générale est de 40 pour 100 000 pour les hommes et de 27 pour 100 000 pour les femmes. En effet, le cancer colorectal est caractérisé par une légère prédominance masculine avec un sexe ratio voisin de 1,5. Au niveau mondial, le cancer colorectal se situe au 3<sup>e</sup> rang des cancers chez l'homme, au 2<sup>e</sup> chez la femme en terme de fréquence.

En France, les taux de mortalité dans la population générale est de 18 pour 100 000 pour les hommes et de 12 pour 100 000 pour les femmes. Entre 1978 et 2000, l'incidence a augmenté d'environ 40% et la mortalité associée a baissé. Cette évolution s'explique par l'amélioration des méthodes de dépistage, de diagnostic précoce et des traitements. En effet, le pourcentage de survie à 5 ans est plus élevé lorsque le cancer est diagnostiqué à un stade très précoce. Il passe de 90% pour un stade très précoce à 35% pour un stade plus avancée et à seulement 5% pour le stade le plus tardif.

## Facteurs de risque

Les cancers colorectaux peuvent être classés en 3 catégories : la forme héréditaire, la forme liée à des maladies inflammatoires et la forme sporadique [1].

La forme héréditaire représente 5 à 10% des cancers colorectaux. Il existe deux principales maladies : la FAP (familial adenomatous polyposis) et la HNPCC (hereditary nonpolyposis colorectal cancer).

La FAP est une maladie autosomale dominante. La plupart des malades présentent une mutation du gène APC (adenomatous polyposis coli). APC promeut la dégradation de la  $\beta$ -caténine et limite la transcription des gènes ciblés par Wnt et impliqués dans la régulation du cycle cellulaire. Parmi ceux-ci se trouve MYH, un important régulateur dans la balance prolifération / apoptose. Il peut également être muté.

La HNPCC est un désordre autosomal dominant, causé par des mutations sur les gènes de réparation de l'ADN comme MLH1, MSH2, MSH3, MSH6, PMS1 et PMS2 [2]. Ces patients ont également un

risque accru de développer des tumeurs dans d'autres organes comme l'estomac, le système biliaire, le pancréas...

En cas de maladies inflammatoires, comme la colite ulcéreuse ou la maladie de Crohn, le risque de déclarer un cancer colorectal augmente avec la durée de la maladie. Il est de 2% durant les 10 premières années, 8% pour 20 années et 18% pour 30 années. Les facteurs de risque pour la forme sporadique sont divers et variés :

- l'âge : le risque est accru après 60 ans et double tous les 10 ans [3].
- le sexe : il existe une légère prédominance masculine.
- l'ablation de la vésicule biliaire (cholécystectomie) serait un facteur de risque
- des facteurs hormonaux : la nulliparité, une grossesse tardive ou une ménopause précoce sont des facteurs de risque
- des facteurs environnementaux. Un régime alimentaire riche en viande et pauvre en fibres, folate et calcium, l'obésité, le diabète, le tabac, l'alcool sont autant de facteurs de risques
- les antécédents familiaux. Le risque de cancer colorectal est multiplié par deux ou trois si un parent de premier ou deuxième degré est affecté et jusqu'à vingt-cinq si plusieurs parents sont touchés. 20% des cas sporadiques ont tout de même un caractère familial sans qu'aucun gène muté connu soit transmis.
- l'histoire personnelle. Un antécédent personnel de polypes colorectaux ou d'une première tumeur multiplie les risques.

## **Classification**

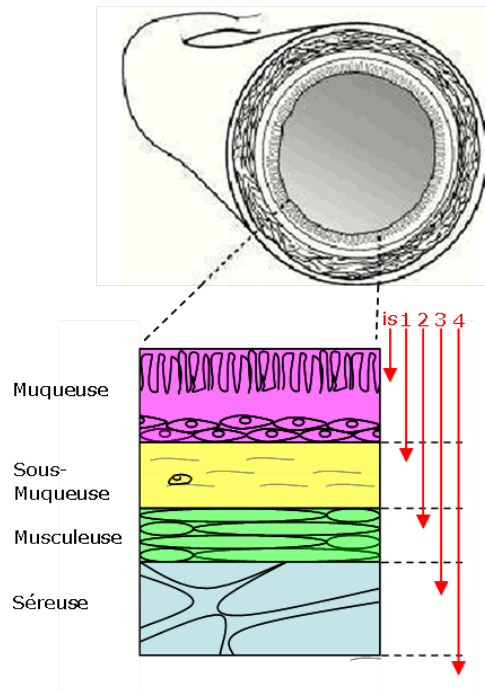
Une fois le cancer colorectal diagnostiqué, le choix de la stratégie thérapeutique se base sur la classification internationale, AJCC/UICC (American Joint Committee on Cancer/Union internationale Contre le Cancer). Cette classification a été originellement décrite en 1932 par C. Dukes [4]. Elle prend en compte l'extension pariétale de la tumeur (stade T), l'existence et le nombre de ganglions lymphatiques envahis (stade N) et la présence ou non de métastases (stade M) (Figure 1). Grâce à ces trois paramètres, les patients sont classés en 4 groupes : Du stade I (une tumeur de petite taille, sans envahissement ganglionnaire ni métastase) au stade IV (tumeur avec métastases) (Figure 1). Cette classification tente de regrouper les différents stades d'extension tumorale de pronostic assez proche. Cependant elle est imparfaite car les groupes de patients sont assez peu homogènes. Ceci est particulièrement vrai pour les cancers de stade II qui regroupent des cancers de gravité variable.

### **Extension pariétale de la tumeur : le stade T**

La maladie débute par une lésion dysplasique de l'épithélium colique. Ainsi, les cancers colorectaux dérivent de la paroi muqueuse. L'extension de ces adénocarcinomes se fait en profondeur atteignant successivement la sous-muqueuse, la musculuse puis la séreuse. L'extension déborde la paroi intestinale et le péritoine, pouvant ainsi former des métastases et envahir les organes avoisinants.

## Stade T

Extension pariétale



## Stade N

Envahissement des ganglions lymphatiques

- N0:** aucun ganglion envahi
- N1:** 1 à 3 ganglions envahis
- N2:** plus de 4 ganglions envahis

## Stade M

Présence de métastases

- M0:** pas de métastase
- M1:** métastases

Stade	T	N	M	Survie globale à 5 ans
I	T1, T2	N0	M0	80-95%
II	T3, T4	N0	M0	68-71%
III	Tout T	N1, N2	M0	30-35%
IV	Tout T	Tout N	M1	0-7%

FIGURE 1 – Classification histologique internationale du cancer du colon

## **Invasion ganglionnaire : le stade N**

Les ganglions sont des structures lymphoïdes souvent situées aux bifurcations des vaisseaux lymphatiques. Ils jouent un rôle important dans la réponse immunitaire. Des cellules tumorales peuvent se détacher de la tumeur primaire et entrer dans des vaisseaux lymphatiques pour envahir les ganglions adjacents de la tumeur. Les cellules tumorales peuvent y être détruites ou se développer en se propageant dans toute la structure ganglionnaire.

## **Métastases à distance : le stade M**

Les métastases sont des amas de cellules tumorales issues de la tumeur primaire et secondairement implantés dans un organe distant. Plusieurs étapes successives sont nécessaires à la survenue d'une métastase : les cellules cancéreuses doivent quitter la tumeur et migrer dans la circulation sanguine ou lymphatique, survivre durant leur trajet, s'arrêter au niveau d'un organe, accomplir une extravasation dans le tissu, puis initier leur croissance et leur vascularisation.

## **Facteurs pronostiques**

Après le diagnostic d'un cancer colorectal, la chirurgie est le traitement de base. Il a pour but l'exérèse du segment colique porteur de la tumeur avec des marges de tissu sain. L'attitude thérapeutique suivant l'acte chirurgical est actuellement guidée par les données histopathologiques de la tumeur. Ainsi, une chimiothérapie adjuvante est le traitement recommandé en cas d'envahissement ganglionnaire (stade III et IV) [1]. Pour les tumeurs localisées (stade I et II), il n'existe pas de standard pour le traitement adjuvant. Or le pronostic des patients de stade II présente des difficultés. En effet, le devenir de ces patients est très hétérogène : près de 30% de ce groupe va faire une récurrence. Une chimiothérapie adjuvante pour l'ensemble des patients présentant un cancer de stade II n'apporterait pas de bénéfice car son efficacité est limitée et expose à une toxicité importante. Chez ceux-ci, les essais d'un tel traitement ont été controversés et une méta-analyse d'essais cliniques prospectifs randomisés de chimiothérapie de ces patients n'a pas montré d'effet bénéfique global sur la survie. Cependant, une autre étude [5] a montré qu'une chimiothérapie adjuvante chez les patients de stades II et III apportait un bénéfice en terme de survie pour certains patients de stade II. C'est pourquoi il est important de mieux identifier les patients à risque pour les faire bénéficier de traitements complémentaires à la chimiothérapie. De nombreuses études s'emploient à chercher de nouveaux marqueurs pronostiques en s'intéressant soit à la tumeur elle-même, soit au micro environnement tumoral.

## **Les facteurs intrinsèques à la tumeur**

L'homéostasie tissulaire est maintenue par un équilibre entre la mort cellulaire et la prolifération. Tout dérèglement de cet équilibre est à l'origine de nombreuses pathologies. Selon le modèle proposé par Vogelstein *et al.* [6], l'apparition de la tumeur se fait après une séquence de mutation. La première étape est associée à la perte d'APC, puis de KRAS, SMAD4 et enfin de TP53. À un point indéterminé, les instabilités chromosomiques et génétiques se développent. C'est pourquoi, afin de

trouver de nouveaux marqueurs prédictifs de rechute ou de sensibilité à un traitement, d'intenses efforts de recherche se sont concentrés sur les facteurs intrinsèques à la tumeur. Plusieurs catégories de marqueurs ont été étudiées :

- Les gènes suppresseurs de tumeur et les oncogènes (APC, K-ras, TP53, SMAD4...).
- APC intervient dans la régulation du cycle cellulaire. Il promeut la dégradation de la  $\beta$ -caténine et limite la transcription des gènes ciblés par Wnt et impliqués dans la régulation du cycle cellulaire. 90% des patients ont une altération de ce gène. Il n'est donc pas surprenant qu'il ne puisse pas servir de marqueur pronostic. La mutation de KRAS n'est démontrée que dans un tiers des cas. Cette mutation compromet la capacité d'hydrolyse du GTP lié à Ras en GDP. Il en résulte un signal permanent stimulant la prolifération cellulaire. Certaines études montrent que cette mutation est de mauvais pronostic [7] tandis que d'autres ne trouvent aucune association [8]. Récemment, l'intérêt de l'évolution de KRAS pour la réponse au traitement par chimiothérapie a été montré [9].

La région où se trouve TP53 (chromosome 17p) perd fréquemment son hétérozygotie. La protéine p53 a un rôle clé lors du dommage de l'ADN. Elle permet la réparation de l'ADN et induit la mort cellulaire au besoin. La mutation de p53 entraîne la dérégulation de la prolifération cellulaire et l'échec des mécanismes apoptotiques, conduisant à une forte instabilité génomique. Ce gène a été étudié à la fois comme facteur prédictif d'une rechute et de sensibilité aux traitements. Cependant les résultats sont variables mais les différences entre les méthodes utilisées ne permettent pas de tirer des conclusions [10].

SMAD4 impliqué dans la voie de signalisation du TGF $\beta$  se situe dans le chromosome 18q, région souvent perdue. Cette délétion de région a été étudiée comme facteur prédictif d'une rechute. Cependant les résultats sont variables : Dans certaines études, elle est associée à un mauvais pronostic [11], dans d'autres elle n'est pas associée à un pronostic particulier [12].

- Les instabilités chromosomique (MSI) et génétique (aneuploïdie) sont respectivement des marqueurs de bon et mauvais pronostic [13].
- Les gènes d'apoptose, Bcl-2 et BAX
- Les gènes de réparation de l'ADN, MSH2 et MLH1
- Les facteurs de croissance et leurs récepteurs,
- Les molécules d'adhésion, CD44, E-cadhérine, ICAM-1
- Les marqueurs de prolifération, Ki-67, Mib-1
- La longueur des télomères
- ...

Cependant, aucun de ces marqueurs n'est utilisé en clinique car ils n'ont pas été validés comme facteurs pronostiques robustes et indépendants [10].

## **Le micro-environnement de la tumeur**

La tumeur se développe dans un micro environnement complexe contenant des fibroblastes, des cellules endothéliales, des facteurs de croissance, des vaisseaux sanguins et lymphatiques, des cellules immunitaires (lymphocytes T, NK, cellules dendritiques, macrophages, ...), des cytokines, des chimiokines. . . De par ses composants, le micro-environnement influence la croissance de la tumeur et sa capacité à faire des métastases. C'est pourquoi ses nombreux composants sont étudiés non seulement dans un but thérapeutique mais aussi pronostique.

Les facteurs de croissance et leurs récepteurs comme  $TGF\alpha$ ,  $TGF\beta$ , HER-2/neu ou EGFR, ainsi que les molécules liées à l'angiogénèse comme VEGF et VEGF-R, n'ont pas été validés en tant que facteurs pronostiques. Toutefois au niveau thérapeutique, la *Food and Drug Administration* (FDA) a approuvé la mise sur le marché d'anticorps ciblant le VEGF et le EGFR [14]. L'étude des cellules immunitaires n'est cependant pas délaissée. En effet, de nombreuses études indiquent que les cellules tumorales colorectales expriment des antigènes (K-ras, p53, CEA,...) qui peuvent induire des réponses immunitaires chez les patients [15]. De plus, des réponses immunitaires anti-tumorales spécifiques ont été mises en évidence [16]. Plusieurs types de cellules immunitaires ont été mises en évidence au site de la tumeur : des cellules du système immunitaire inné (comme les cellules dendritiques [17]) et du système immunitaire adaptatif (comme les T helper 1 ou Th1[18] [19]). D'une manière générale, la présence de lymphocytes infiltrant la tumeur (TILs) joue un rôle dans l'inhibition de la croissance des tumeurs et est associée à un meilleur pronostic [20][21].



# Immunité anti-tumorale

## Lymphocytes T et cancer

### Quelques généralités

La réaction immunitaire anti-tumorale est complexe. Elle est composée de différents types de cellules infiltrantes et sont trouvées dans différentes régions de la tumeur : au sein même de la tumeur, au front de progression de la tumeur ou dans des îlots lymphoïdes à distance de la tumeur. Une évaluation du statut immunitaire doit prendre en compte la densité, la nature et leurs capacités à détruire les cellules tumorales.

Si le micro-environnement peut influencer le devenir de la tumeur, la tumeur influence elle-aussi le micro-environnement et la réponse immunitaire en particulier. En effet, au contact de la tumeur, les lymphocytes T (LT) entreraient dans un état de tolérance fonctionnelle : ils ne contiennent pas de perforine et ne produisent pas d'IFN $\gamma$  après re-stimulation [22]. Les cellules tumorales développent donc des mécanismes d'échappement au système immunitaire.

### Cytotoxicité

La capacité des cellules immunitaires à tuer les cellules tumorales est donc très importante. L'activité cytotoxique est une des réponses du système immunitaire. Elle permet de détruire spécifiquement et par contact direct les cellules exprimant des antigènes cibles. Trois types de cellules ont cette capacité : les macrophages, les cellules NK (natural killer) et les lymphocytes T cytotoxiques. Les macrophages ont la capacité de libérer dans le phagolysosome, les formes réactives de l'oxygène et des enzymes lysosomiales. Les NK et les lymphocytes T cytotoxiques utilisent la même machinerie lytique : ils libèrent des molécules cytotoxiques au contact de la cellule. Ces molécules sont responsables directement ou indirectement de la mort cellulaire : soit par un mécanisme de nécrose secondaire à des lésions membranaires, soit par un processus apparenté à l'apoptose. Il existe deux grandes catégories de molécules cytotoxiques : celles qui sont stockées dans des granules à l'intérieur de la cellule en attendant d'être relarguées, et celles qui sont à la membrane ou directement sécrétées.

La perforine, la granulysine et les granzymes sont synthétisés et stockés sous forme inactive. La perforine, après relargage, se polymérise, s'insère dans la membrane de la cellule cible et forme des pores membranaires. Ces pores permettent les flux d'eau entre l'extérieur et l'intérieur de la cellule ainsi que le passage de petites molécules. À forte concentration, elle est responsable de la lyse osmotique, à plus faible concentration, elle permet le passage des granzymes dans le cytoplasme de

la cellule cible.

Il existe plusieurs type de granzymes : de granzyme A à H. Ce sont toutes des protéases. Le granzyme B est le plus important dans l'activité cytotoxique. Il clive les précurseurs des caspases et déclenche une cascade de réactions enzymatiques aboutissant à l'apoptose.

La granulysine est une protéine qui interagit avec les lipides membranaires. Elle active les enzymes responsables de leur dégradation. La concentration cellulaire en céramide, produit de la dégradation des lipides, active la cascade des caspases, ce qui aboutit à l'apoptose des cellules cibles.

Les cellules cytotoxiques expriment à leurs surfaces FasL (ou CD178), le ligand de Fas. Fas est une molécule membranaire de la famille du TNF (tumor necrosis factor) possédant un domaine cytoplasmique de mort. L'interaction Fas/FasL dans les cellules exprimant le Fas (ou CD95) à leurs surfaces, aboutit à l'apoptose. La lymphotoxine utilise un mécanisme analogue. C'est une cytokine qui a deux formes : La forme sécrétée (lymphotoxine  $\alpha$ ) et une forme membranaire (lymphotoxine  $\beta$ ). Les deux formes se lient à des récepteurs de la famille des TNF, ce qui aboutit à l'apoptose de la cellule.

L'activité cytotoxique des cellules implique que ces cellules soient activées et qu'elles aient eu une interaction suffisante entre le TCR et le peptide [23]. En effet, les cellules immunitaires doivent avoir reconnu les cellules tumorales comme du non-soi ou du soi modifié pour s'activer se différencier en effecteur et les détruire. Cette activation est possible car les cellules tumorales présentent des antigènes tumoraux.

### **Antigènes tumoraux et activation des lymphocytes**

Plusieurs familles d'antigènes de tumeur peuvent être distinguées : les antigènes résultant de gènes viraux, les antigènes spécifiques des tumeurs, les antigènes de différenciation spécifiques du tissu, les antigènes résultant de mutations, les antigènes résultant de gènes amplifiés et sur-exprimés et les antigènes de protéines anormalement modifiées post-transcriptionnellement [14].

Certains virus jouent un rôle initiateur de cancérogenèse. C'est le cas du virus d'Epstein-Barr dans le cas du cancer du nasopharynx, du virus de l'hépatite B dans l'hépatocarcinome ou du virus du papillome dans le cancer du col utérin. Des gènes viraux sont à l'origine de protéines cellulaires anormales et antigéniques.

La plupart des antigènes spécifiques des tumeurs sont codés par des gènes non exprimés dans les tissus adultes normaux excepté les cellules germinales males. Or celles-ci n'expriment pas de molécules CMH (complexe majeur d'histocompatibilité) et donc ne sont pas capables de présenter ces antigènes. C'est le cas des familles des protéines MAGE, BAGE, et CAGE [24] [25].

Les tumeurs expriment des antigènes spécifiques d'un tissu. Ainsi les cellules du mélanome expriment les gènes comme la tyrosinase, Melan-A/MART-1, Pmel17/gp100, TRP-1/gp75 ou TRP-2, qui sont exprimés également par les mélanocytes.

Une mutation peut conduire à une protéine anormale. Par exemple, les mutations de KRAS et p53 présentes dans certains cancers colorectaux, produisent des protéines anormales reconnues par le système immunitaire [15].

Des protéines sur-exprimées peuvent être reconnues comme des antigènes tumoraux. Ces protéines ne sont pas exprimées dans les tissus sains à l'exception des testicules et du placenta, mais sont présentes dans les cellules du mélanome. Elles peuvent activer des lymphocytes T [26][27]. Les antigènes carcinoembryonnaires ACE, Her-2/neu entrent aussi dans cette catégorie.

Des protéines qui sont modifiées anormalement au niveau post-transcriptionnel peuvent être reconnues comme antigènes tumoraux. C'est le cas de la cycline B1. Cette protéine se trouve normalement au niveau du noyau et contrôle la transition G2-M du cycle cellulaire. Dans beaucoup de cellules tumorales, elle est constitutivement sur-exprimée au niveau du cytoplasme. Le système immunitaire des patients avec des tumeurs exprimant anormalement la cycline B1 peuvent la reconnaître comme antigène [28].

Des anomalies d'enzyme de réparation de l'ADN telles que MSH1, MSH2, MSH3, MSH6, PMS1 et PMS2 sont fréquentes dans les cancers et en particulier les cancers colorectaux. Ainsi 15% des patients atteints de cancer colorectaux présentent des instabilités des microsattelites (MSI) associés à ces anomalies. Comme les autres cellules de l'organisme, les cellules tumorales exposent les peptides par l'intermédiaire des molécules du CMH (complexe majeur d'histocompatibilité) de classe I. A cause de ces défauts d'enzyme de réparation de l'ADN, ces patients présentent de nombreuses autres protéines mutées. Des lymphocytes spécifiques de ces peptides mutés ont été trouvés. Ces lymphocytes T sont capables de lyser les cellules présentant ces protéines [29][30][31][32][33].

Bien que reconnues par les lymphocytes, les cellules tumorales sont néanmoins des cellules du soi dont la fonction présentatrice d'antigènes est peu efficace. Le rôle des cellules présentatrices d'antigènes (CPA) est essentiel pour initier une réponse immunitaire antitumorale. Ces cellules ont des capacités de phagocytose et d'endocytose qui leur permettent d'internaliser des antigènes exogènes. Elles ont la capacité de les présenter ensuite sur des molécules du CMH de classe I et II. Une fois chargées, elles migrent dans un ganglion de drainage pour initier la réponse cellulaire T : L'antigène présenté par le CMH de classe II ou I est alors reconnu par un lymphocyte T spécifique de cet antigène, respectivement CD4+ ou CD8+. Ces lymphocytes ne sont activés que si le premier signal de reconnaissance du peptide est accompagné de signaux dits de co-stimulation. C'est le cas si la CPA a reçu des signaux de danger et si elle est mature. Les lymphocytes T CD8 vont recevoir leur signal de co-stimulation par l'interaction du CD28 avec B7 situé sur la CPA et les lymphocytes T CD4 par l'interaction du CD40L avec le CD40 situé sur la CPA. Un lymphocyte T CD8 peut s'activer sans signaux de co-stimulation si un lymphocyte T CD4 "helper" par ailleurs déjà activé, libère des facteurs de croissance. Activés, les lymphocytes T CD8 se différencient en cellules effectrices, acquérant ainsi leurs capacités cytotoxiques. Il existe différents types de lymphocyte T CD4 "helper" ayant différents rôles et qui vont orienter différemment la réponse immunitaire. Les lymphocytes Th1 (T helper 1) orientent vers une réponse cellulaire, les Th2 (T helper 2) orientent vers une réponse humorale, les Treg (T régulateurs) induisent une tolérance et diminuent la réaction immunitaire et les Th17 (T helper 17) induisent l'inflammation. Il existe aussi les TFH (T follicular helper) qui sont spécialisés dans le dialogue avec les lymphocytes B [34]. Cependant il existe encore une discussion sur la classification précise de cette dernière catégorie.

## **Les voies de différenciation des lymphocytes T**

### **Les Lymphocytes T helper 1**

Les Th1 sont, avec les Th2, les deux premières populations de lymphocytes T helper décrites. En effet, en 1986, Mossmann *et al.* décrit deux populations distinctes parmi les lymphocytes T CD4+. Ces deux populations ont des profils de sécrétion et des fonctions différentes et distinctes [35]. Les Th1

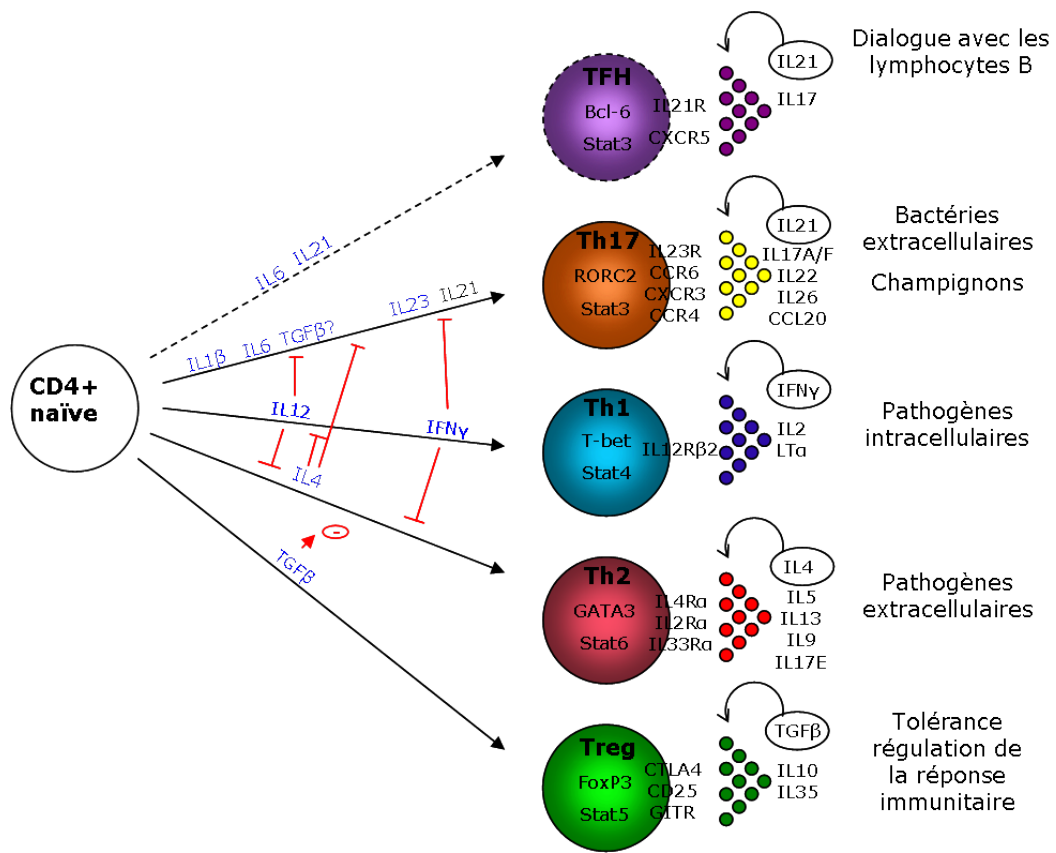


FIGURE 2 – Les différentes voies de différenciation des T helper

sont importants pour la réponse immunitaire contre les pathogènes intracellulaires. Ils se différencient en présence d'interleukine 12 (IL-12) et interféron- $\gamma$  (IFN $\gamma$ ). Les Th1 sécrètent de l'IFN $\gamma$  formant une boucle de retro-contrôle positive, mais aussi de l'IL2 et de la lymphotoxine  $\alpha$ . L'IFN $\gamma$  inhibe également la différenciation des lymphocytes T en Th2 ou Th17 (Figure 2). Stat4 est un facteur de transcription spécifique des Th1, se situant au bout de la chaîne de transduction initié par l'interaction IL-12/IL12R. T-bet, un autre facteur de transcription spécifique, contrôle la transcription de l'IFN $\gamma$ . Les Th1 forment une boucle de retro-contrôle positive. En effet, l'IFN $\gamma$  active les macrophages, promeut l'expression de molécules d'adhésion pour la migration des leucocytes, l'expression du CMH de classe I sur toutes les cellules et l'activité cytotoxique des cellules immunitaires.

## Les Lymphocytes T helper 2

La deuxième population de lymphocytes T helper initialement décrites par Mossmann *et al.* [35] s'appelle les Th2. Ces lymphocytes sont importants pour la réponse immunitaire contre les parasites extracellulaires et favorisent une réponse humorale. Ils se différencient en présence d'IL-4. Les Th2, une fois différenciés, sécrètent de l'IL-4, formant ainsi une boucle de retro-contrôle positive. Ils sécrètent également de l'IL-5, l'IL-13, l'IL-9, l'IL-17E et de l'amphiréguline (Figure 2). L'IL-4 permet également le switch vers les Immunoglobulines de classe E (IgE) et G1 (IgG1) des lymphocytes B. Via leur récepteur sur les basophiles et les mastocytes, les IgE permettent la libération de l'histamine, la sérotonine ou encore le TNF $\alpha$ . L'IL-5 permet de recruter les éosinophiles. L'IL-13 est une cytokine effectrice pour l'expulsion des helminthes. L'IL-9 permet le développement des mastocytes et basophiles et la production de mucine par les cellules épithéliales. L'amphiréguline est un membre de la famille des EGF (epidermal growth factor). Elle induit la prolifération des cellules épithéliales. L'IL-17E (ou IL-25) amplifie la production des interleukines 4, 5 et 13. Il induit également la production des chimiokines CCL5 et CCL11 qui recrutent les éosinophiles. [36] L'IL-4 active le facteur de transcription spécifique des Th2, Stat6, qui va à son tour activer un autre facteur de transcription spécifique, GATA3. Ce dernier contrôle la transcription des cytokines spécifiques des Th2, IL-4, IL-5 et IL-13.

## Les Lymphocytes T régulateurs

En 1994, Chen *et al.* ont montré qu'un régime particulier induit une tolérance orale grâce à des lymphocytes T CD4 produisant du TGF $\beta$  [37]. Cette population a été nommée Th3. Aujourd'hui, les lymphocytes T régulateurs sont composés de plusieurs sous-groupes qui se ressemblent phénotypiquement : les nTreg (natural T regulatory) et les iTreg (induced T regulatory). Hormis les Th3 qui produisent le TGF $\beta$ , il existe les TR1 qui produisent de l'IL-10. Cependant les TR1 ne formeraient pas une lignée distincte mais représenteraient plutôt un état de chaque lignée : Tous les lymphocytes T CD4 (Th1, Th2, Th17) sont capables de produire de l'IL-10 dans certaines circonstances [36]. La plupart des lymphocytes T régulateurs expriment le facteur de transcription FoxP3 qui est induit par le TGF $\beta$  (Figure 2). Les Treg induisent une tolérance et diminuent la réaction immunitaire grâce aux interleukines 10 et 35 et le TGF $\beta$ .

## Les Lymphocytes T helper 17

Les Th17 sont la dernière lignée à avoir été mise en évidence. Elle a été découverte dans le cadre d'une maladie, l'encéphalomyélite auto-immune expérimentale (EAE) [38]. Elle a été définie comme la population lymphocytaire CD4 produisant de l'IL-17. Ils sont importants pour la réponse immunitaire contre les bactéries et les champignons. Les Th17 ont pour facteur de transcription spécifique RORC2 (équivalent de ROR $\gamma$ t chez la souris), qui permet la transcription de l'IL-17. Outre l'IL-17A et de l'IL-17F, les Th17 sécrètent de l'IL-21, l'IL-22 et l'IL-26 (Figure 2).

L'IL-17A induit des cytokines inflammatoires, comme IL-6 ou IL-8. L'IL-17A et F recrutent et activent les neutrophiles. L'IL-21 produit sert à la boucle de retro-contrôle positive en promouvant l'expression du récepteur à l'IL-23. Elle intervient également dans la différenciation des lymphocytes en TFH. Le rôle précis de l'IL-22 n'est pas connu, mais elle intervient dans la protection contre les bactéries [39]. Son expression est inhibée par la présence de TGF $\beta$ . L'IL-26 fait partie de la même famille que l'IL-22. Cependant son rôle n'est pas connu.

Les conditions de différenciation en Th17 sont mal connues et controversées. Chez la souris, IL-6 et le TGF $\beta$  sont nécessaires pour la différenciation puis l'IL-21 et l'IL-23 pour l'amplification et le maintien du phénotype. Chez l'Homme, l'IL-6 et l'IL-1 $\beta$  sont nécessaires pour la différenciation puis comme chez la souris, l'IL-21 et l'IL-23 pour l'amplification et le maintien du phénotype. La controverse concerne le TGF $\beta$  : Quelques études ont montré que le TGF $\beta$  n'était pas nécessaire [40] [41], tandis que d'autres ont montré qu'il est indispensable [42] [43] [44]. Selon ces dernières études, la quantité de TGF $\beta$  nécessaire à la différenciation en Th17 est très faible sinon les cellules se différencient en Treg. Une petite contamination par du sérum ou des cellules qui en produisent suffit à apporter la quantité nécessaire.

## Les lymphocytes T Follicular Helper

Les TFH sont des lymphocytes T spécialisés dans le dialogue avec les lymphocytes B. Ces cellules se situent dans les centres germinatifs, au contact des lymphocytes B. Il existe à ce jour deux hypothèses pour leur classification : Des données suggèrent plutôt que tous les types de lymphocytes T helper peuvent développer les fonctions des TFH lorsqu'ils se trouvent dans un follicule. Mais d'autres données, récentes, suggèrent que ces cellules forment une classe à part avec des conditions de différenciation et un phénotype distincts des autres lymphocytes T helper [45].

Les TFH comme les lymphocytes B, expriment fortement le récepteur de chimiokine CXCR5 ce qui leur permet de migrer vers les zones des follicules riches en CXCL13. Outre CXCR5, les TFH expriment spécifiquement le facteur de transcription Bcl-6 [45], la chimiokine CXCL13, le récepteur à l'IL-21, l'IL-17 et l'IL-21. L'IL-21 et l'IL-6, sont indispensables à leur différenciation, formant la boucle de retro-contrôle positive (Figure 2).

Les TFH et les lymphocytes B interagissent grâce à une série de molécules : TCR, CD40L, ICOS, SLAM et OX40 présents sur les TFH interagissent avec respectivement CMH de classe II, CD40, ICOSL, SLAM et OX40L. L'environnement cytokinique pro-Th1, Th2 ou T régulateur permet un switch vers un type particulier d'immunoglobuline. Ainsi la présence d'Il-4 favorise les IgG1 et les IgE, l'IFN $\gamma$  les IgG2a, l'IL-10 et le TGF $\beta$  les IgA [45].

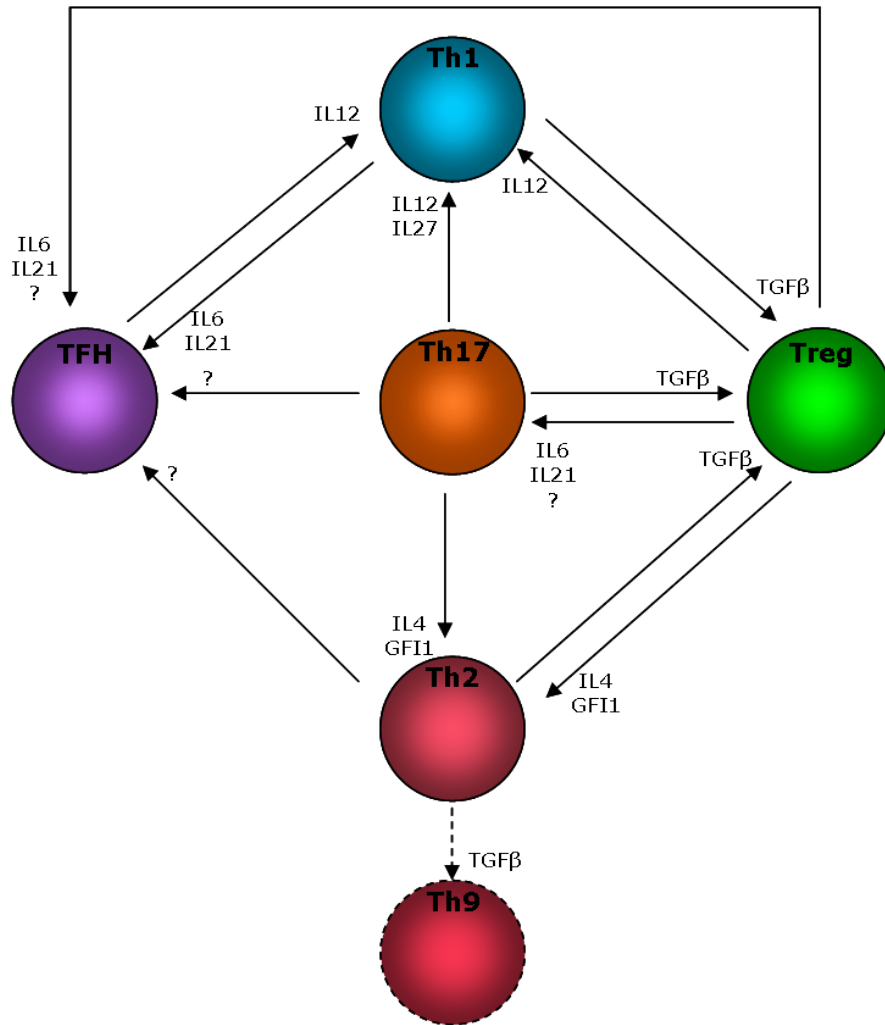


FIGURE 3 – Les conversions observées des T helper

### La plasticité des lymphocytes T Helper

La différenciation des lymphocytes T dans une voie n'est pas irréversible. Ces cellules font preuve de plasticité en acquérant, dans certaines conditions, des caractéristiques d'une autre voie (Figure 3). Ainsi, en présence de  $TGF\beta$ , les Th1, Th2 et Th17 deviennent capables de produire de l'IL-10 [46]. Les Th2, dans ce cas, sont également capables de produire de l'IL-9 mais plus d'IL-4, IL-5 ou IL-13. Elles sont alors nommées Th9 [47].

Des cellules co-exprimant le facteur de transcription FOXP3, spécifique des T régulateurs, et RORC2, spécifique des Th17 ont été identifié [48]. Elles sont capables de produire de l'IL-17. Des lymphocytes T nommés Th1/17 ont été décrits [49]. Elles expriment les facteurs de transcription T-bet et RORC2. Elles produisent en même temps de l'IFN $\gamma$  et de l'IL-17. Les Th17 semblent également devenir des Th2 en présence d'IL-4 et GFI-1 (Growth factor independent 1) [50].

Les TFH et les Th17 se différencient dans un environnement cytokinique proche. Toutes deux ont besoin d'IL-6 et d'IL-21. Le passage Th17 vers TFH a été observé, bien que les conditions ne soient pas encore clairement établies [45]. Récemment il a été montré que dans les plaques de Peyer de

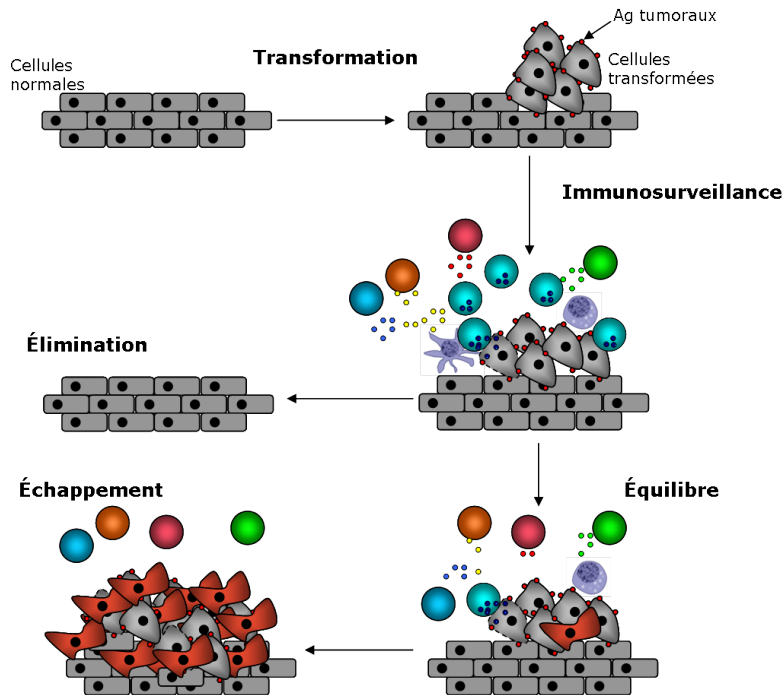


FIGURE 4 – Les phases de l’immunoediting

la souris, les TFH proviennent préférentiellement de lymphocytes T régulateurs FOXP3+. Les TFH peuvent également provenir de Th2 en réponse à des antigènes d’helminthes [51].

## Modèle d’équilibre/échappement de la tumeur au système immunitaire

En 1970, Burnet a proposé ce concept d’immunosurveillance [52] : Le système immunitaire pourrait protéger contre le développement des cancers d’origine non virale. Aujourd’hui de nombreuses données, chez la souris comme chez l’Homme, montrent l’importance du système immunitaire inné et adaptatif. En effet, les patients immunodéprimés (comme dans le cas du SIDA ou de transplantation d’organes) ont une fréquence plus élevée de cancers [53]. Toutefois, la présence d’une tumeur cliniquement détectable suggère l’échappement de la tumeur. C’est le concept de l’immunoediting. Il existe trois phases : élimination, équilibre et échappement [54].

La première phase, la phase d’élimination ou d’immunosurveillance, consiste à la reconnaissance et à l’élimination des cellules transformées par le système immunitaire inné et adaptatif. La deuxième phase, l’équilibre ou la phase de persistance, est une période où la tumeur persiste mais son expansion est empêchée par la pression immunitaire. La troisième phase, l’échappement, commence lorsque la balance entre le système immunitaire et la croissance tumorale penche en faveur de la tumeur. Elle est le résultat de l’apparition de cellules tumorales capables de résister à la pression immunitaire. Une tumeur cliniquement décelable grossit alors progressivement (Figure 4).



## La surveillance du système immunitaire inné

L'initiation de la réponse immunitaire débute lorsque le système immunitaire inné est alerté. Smyth *et al.* proposent que ce soit la dislocation de l'intégrité du tissu qui induit la sécrétion de molécules pro inflammatoires [55]. En effet, le développement d'une tumeur déclenche un remodelage du stroma, des mécanismes d'angiogénèse, une invasion tissulaire. Les signaux inflammatoires et les chimiokines peuvent être produites par la tumeur elle-même [56].

Selon Matzinger, les cellules dendritiques agissent comme des sentinelles en surveillant le stress tissulaire, les éventuels transformations ou dommages et les signaux de danger comme les HSP (heat shock proteins) (relâchées en cas de dommage cellulaire ou de nécrose), des produits de la dégradation de la matrice extracellulaire ou des facteurs pro inflammatoires [57]. Un de ces signaux permet l'activation et la maturation des cellules dendritiques qui le perçoivent, initiant ainsi la réponse immunitaire.

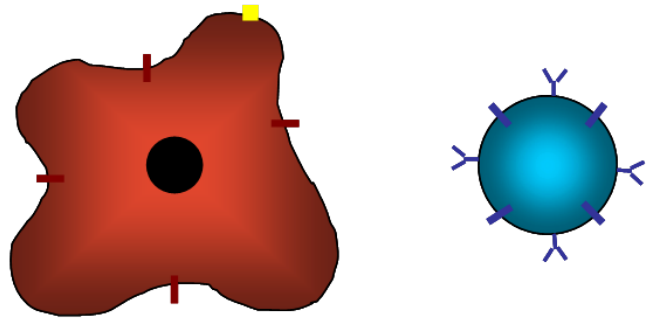
Les cellules NK (natural killer), les macrophages, les lymphocytes T  $\gamma\delta$ , ou les NKT (natural killer T) peuvent être directement recrutés au site de la tumeur. Ces cellules peuvent reconnaître des molécules tumorales par l'intermédiaire de leurs récepteurs. Elles ont une activité cytotoxique et sécrètent des cytokines qui amplifient la réponse immunitaire, notamment de l'IFN $\gamma$  et de l'IL-12. Bien que peu étudiées, les cellules de l'immunité inné sont importantes pour une réponse immunitaire anti-tumorale. Les NK, NKT, les lymphocytes T  $\gamma\delta$  et les neutrophiles sont nécessaires. En effet, les souris déficientes sont plus susceptibles de développer des tumeurs [58] [59]. Chez l'Homme, les patients atteints du syndrome de Chediak-Higashi ont un risque 200 fois plus élevé de développer des tumeurs. Ces malades ont une mutation dans le gène CHS-1 (Chediak-Higashi syndrome 1) qui intervient dans la sécrétion des lysosomes. Cela a pour conséquence une fonction anormale des cellules NK [60]. La perforine, impliquée dans la cytotoxicité, peut aussi être sujette aux mutations. Ces patients, dont la perforine n'est plus fonctionnelle, ont des prédispositions aux lymphomes [61]. Les NK sont aussi importants chez l'Homme. En effet, des combinaisons particulières d'allèles HLA et de KIRs (Killer-cell Immunoglobulin-like Receptors) prédisposent à des cancers cervicaux [62].

## La surveillance du système immunitaire adaptatif

Les cellules présentatrices d'antigènes, recrutées et activées au site tumoral, vont migrer vers les ganglions lymphatiques où elles vont induire l'activation des lymphocytes naïfs spécifiques des antigènes tumoraux. Les lymphocytes T CD4, une fois activés et différenciés en Th1, vont faciliter l'activation des lymphocytes T CD8. Les lymphocytes T, CD4 et CD8, spécifiques des antigènes tumoraux, migrent au site tumoral où ils participent au rejet de la tumeur. Les lymphocytes T CD4 vont produire des cytokines et un micro-environnement favorable à la réponse immunitaire, tandis que les lymphocytes T CD8 vont avoir une activité cytotoxique. L'importance des cellules de l'immunité adaptative a été prouvé chez les souris déficientes : Elles développent des tumeurs [55]. Chez l'Homme, des corrélations entre la présence de lymphocytes T, et notamment une réponse de type Th1, et une meilleure survie des patients ont été observées [18] [19]. Les réponses immunitaires ont la capacité d'éliminer complètement la tumeur en développement (figure 4). Mais le rejet des tumeurs est un processus continu : Il doit être répété à chaque apparition de nouvelles cellules neoplasiques, avec une antigénicité variable. Toutefois, des patients développent des tumeurs, témoignant de l'échec de réponses immunitaires et de l'échappement de tumeurs.

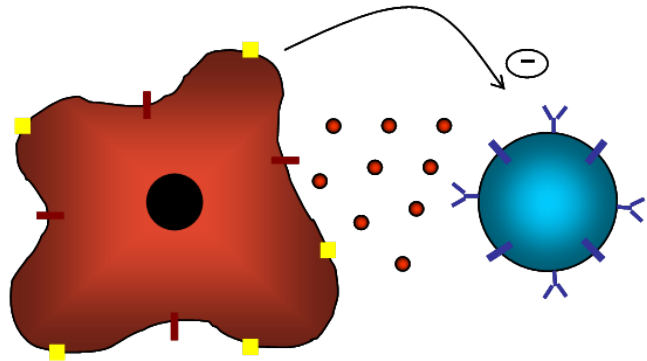
### Ignorance

Le système immunitaire ignore la tumeur.  
Les cellules tumorales présentent peu d'antigènes à leurs surfaces.



### Tolérance

Le système immunitaire tolère la tumeur.  
Les cellules tumorales induit la tolérance du système immunitaire par immunosuppression.



### Résistance à l'apoptose

Le système immunitaire tente de détruire la tumeur.  
Les cellules tumorales résistent à la lyse par le système immunitaire.

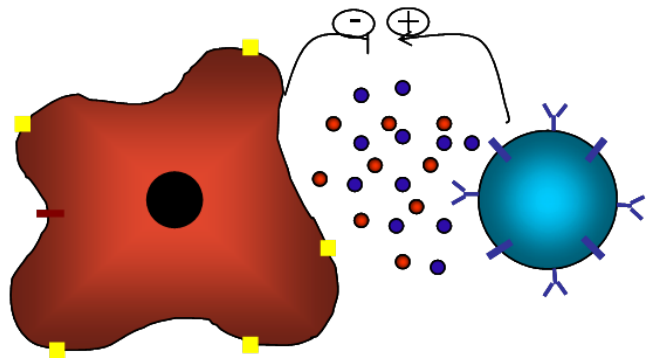


FIGURE 5 – Les mécanismes d'échappement de la tumeur au système immunitaire

## Mécanismes d'échappement de la tumeur au système immunitaire

Le processus qui conduit à l'apparition d'une tumeur est long et progressif. Durant tout ce temps, les cellules tumorales subissent la pression du système immunitaire. Elles développent ainsi progressivement différentes stratégies leur permettant d'échapper au système immunitaire. On peut distinguer trois différentes stratégies : L'échappement par ignorance, par tolérance et par résistance. Il y a également un effet négatif du système immunitaire qui peut promouvoir la progression tumorale.

### Échappement immun par ignorance

Malgré l'existence d'antigènes tumoraux, les cellules tumorales sont souvent caractérisées par une très faible immunogénicité (figure 5). Comme la tumorigénèse est un processus lent, la tumeur peut

progresser pendant des années sans déclencher de signaux de danger. Ainsi le nombre restreint de cellules dans les tumeurs de petites tailles ne doit pas permettre de déclencher une réponse immunitaire. De plus, l'absence de "phase aiguë d'infection" peut contribuer à une faible activation précoce du système immunitaire inné. Par conséquent, le déclenchement d'une réponse immunitaire adaptative dirigée contre la tumeur peut être retardé et altéré. Cette ignorance est donc passive.

Pour échapper à la réponse immunitaire médiée par les lymphocytes T, les cellules tumorales expriment plus faiblement le CMH de classe I. Elles peuvent aussi diminuer l'expression des molécules impliquées dans l'apprêtement des antigènes et dans leur présentation par le CMH de classe I (figure 5). Ainsi, les expressions de TAP1 (transporter associated with antigen processing 1), LMP2 (low-molecular-mass protein 2), LMP7 et la tapasine sont progressivement perdues lors du développement des carcinomes colorectaux [63]. Certaines tumeurs peuvent également transférer les antigènes tumoraux de la surface membranaire dans le cytoplasme. Les antigènes sont alors inaccessibles et donc invisibles au système immunitaire. Il a été également suggéré que les cellules cancéreuses antigéniques soient enfouies au cœur de la tumeur solide. D'autre part, le stroma de la tumeur pourrait empêcher le relargage efficace des antigènes tumoraux [64]. Les cellules tumorales peuvent également induire la formation de la sialomucine, une mucoprotéine qui entoure et masque les antigènes tumoraux de surface [65]. La sialomucine est un composant normal du soi, ce qui empêche toute activation du système immunitaire.

### **Échappement immun par tolérance**

Lorsque la tumeur n'est plus capable d'esquiver sa détection par le système immunitaire, elle peut induire une tolérance de celui-ci (figure 5). Elle peut agir sur les cellules présentatrices d'antigènes ou sur les lymphocytes T. C'est un mécanisme actif.

La tumeur peut mettre en place des mécanismes qui vont empêcher l'initiation de la réponse immunitaire. Ils visent essentiellement à réduire le potentiel stimulateur des cellules présentatrices d'antigènes. Ainsi de nombreux mécanismes empêchent la maturation des cellules dendritiques. En effet, les cellules dendritiques immatures sont immunosuppressives tandis que matures, elles deviennent activatrices. De nombreuses études montrent que la différenciation et la maturation de ces cellules sont altérées par de nombreux facteurs présents dans le micro environnement tumoral. La plupart de ces facteurs sont sécrétés par la tumeur elle-même. L'IL-6 et le M-CSF (macrophage colony-stimulating factor) produits par certaines tumeurs inhibent l'activité de présentation antigénique des cellules dendritiques en induisant la différenciation de leurs précurseur en macrophages [66]. Le VEGF est connu pour son rôle dans l'angiogénèse mais il inhibe également la différenciation et la maturation des cellules dendritiques [67]. Un grand nombre de tumeurs humaines sont capables de produire de fortes quantités de cyclooxygénase-2 (COX2) qui induit à son tour la production de la prostaglandine E2 (PTGS2) dans le microenvironnement tumoral [68]. La PTGS2 est capable de bloquer la différenciation et les fonctions des cellules dendritiques en diminuant notamment l'expression des molécules du CMH de classe I et II [69][70]. Les gangliosides sécrétés par certaines tumeurs peuvent également bloquer la différenciation des DC chez l'homme [71]. Enfin, l'activité des cellules dendritiques peut être altérée sous l'action immunosuppressive des cytokines IL-10 et TGF $\beta$  pouvant être sécrétés par les tumeurs. L'IL-10 est capable de réduire la cytolyse induite par les lymphocytes T antitumoraux en inhibant, chez les cellules dendritiques, l'expression des molécules du CMH de classe I et II, des molécules de transport associée à l'apprêtement des antigènes TAP1 et TAP2, ainsi que des molécules

d'adhésion intercellulaire ICAM-1 et de costimulation B7.1 et B7.2. L'IL-10 est également impliquée dans l'apoptose des cellules dendritiques, soit directe, soit causée par l'intermédiaire des cellules NK. Comme l'IL-10, le TGF $\beta$  est impliqué dans l'inhibition de la maturation des cellules dendritiques. L'IL-10 altère également la qualité de la réponse immunitaire adaptative. En effet, l'IL-10 inhibe la production de l'IL-12 par les cellules dendritiques et les monocytes et de l'IFN $\gamma$  par les cellules NK. L'établissement d'une réponse de type Th1 est bloquée, déviant ainsi vers une réponse de type Th2.

La tumeur peut également agir sur le système immunitaire adaptatif. Elle vise à inhiber les fonctions effectrices des lymphocytes T. Les cytokines IL-10 et TGF $\beta$  sont toutes deux impliquées dans la génération des lymphocytes T régulateurs. Les Treg ainsi formés, produisent à leur tour ces deux cytokines, soutenant et amplifiant l'environnement immunosuppresseur de la tumeur. Ils expriment aussi à leur membrane CTLA-4 molécule inhibitrice qui interagit avec les molécules de co-stimulation B7 des cellules dendritiques. Immunosuppressives, celles-ci secrètent de l'IDO (indoléamine-2,3 dioxygénase) qui bloque la progression du cycle cellulaire des lymphocytes T. Les cellules tumorales sont également capables de produire de l'IDO [72]. Le TGF $\beta$  agit directement sur les lymphocytes T : Il induit la diminution de l'expression des gènes clés de leurs fonctions effectrices, comme la perforine, les granzymes A et B ou l'IFN $\gamma$ . Les cellules tumorales peuvent interagir directement avec les lymphocytes T grâce à des récepteurs membranaires. Ainsi, la molécule de costimulation négative B7-H1 exprimé par de nombreuses cellules tumorales, est impliquée dans la mort des lymphocytes T effecteurs au site tumoral [73]. En effet, son ligand PD-1 est plus exprimé par les lymphocytes T effecteurs que par ceux en phase précoce d'activation. Ceci pourrait expliquer que les signaux induits par PD-1 induisent l'anergie des Lymphocytes T naïfs lors de la présentation des antigènes par les cellules présentatrices d'antigènes mais induisent la mort cellulaire des lymphocytes T effecteurs au contact de la tumeur. Un autre récepteur RCAS1, peut être également exprimé par certaines tumeurs et a été impliqué dans l'apoptose des lymphocytes T effecteurs antitumoraux [74].

### **Échappement immunitaire par résistance à l'apoptose**

Les cellules tumorales développent également des mécanismes leur permettant de résister à la lyse par le système immunitaire (figure 5). La sur-expression d'inhibiteurs ou de molécules anti-apoptotiques et des récepteurs leurres sont des mécanismes utilisés par les cellules tumorales. Par exemple, la sur-expression de la PI9 (serin-Protease Inhibitor 9) bloque efficacement l'apoptose induit par la perforine et les granzymes B [75]. Les récepteurs de mort, comme FasL ou la caspase 8 ainsi que les molécules intervenant dans la voie apoptotique comme p53 [76], peuvent être moins exprimés ou muté à la surface de la cellules tumorales. Un inhibiteur de la caspase 8, FLIP (FLICE-like inhibitory protein), peut être sur-exprimé. Les cellules tumorales peuvent aussi produire DCR3, un récepteur soluble liant et neutralisant FasL. Il existe aussi des récepteurs leurres TRAIL (TNF-related apoptosis-inducing ligand) [53]. Les cellules tumorales peuvent aussi éviter l'apoptose en sur-exprimant des molécules anti-apoptotiques comme Bcl-2.

### **Effets promoteurs du système immunitaire sur la progression tumorale**

La tumeur se développe en étant en interaction avec le système immunitaire. Elle a développé des mécanismes lui permettant d'échapper au système immunitaire et à sa lyse. Elle est également ca-

pable d'utiliser le système immunitaire et ses molécules pour sa progression. Ainsi, dans des modèles murins, les lymphocytes T CD4 contribuent à la carcinogénèse [77].

Les cellules tumorales bloquent les réponses immunes de type Th1, déviant ainsi la réponse vers une réponse de type Th2. Or IL-4, une cytokine de type Th2, peut induire une sur-expression de molécules antiapoptotiques comme FLIP et BCL-XL par la tumeur. Ce mécanisme a été montré dans des cancers humains de sein, prostate et vessie [78]. La tumeur attire des cellules immunitaires immatures ce qui crée un environnement immunosuppresseur : Des antigènes tumoraux comme gp100, MUC1 ou CEA attirent les cellules dendritiques. En effet, la gp100 se fixe sur un récepteur de chimiokine, CCR2 [79]. Toutefois les antigènes tumoraux n'induisent pas la maturation des cellules dendritiques. Les médiateurs pro-inflammatoires comme la  $\beta$ -defensine ou la chimiokine CXCL12, sécrétés par le système immunitaire, recrutent au site tumoral des cellules dendritiques plasmacitoïdes (pDC) et des cellules vasculaires. Or ces deux types cellulaires contribuent à l'angiogénèse : Les VLC en se différenciant en cellules endothéliales en présence de VEGF et les pDC en sécrétant du TNF et CXCL8 (ou IL-8) cytokines pro-angiogéniques [80]. L'IL-8 induit également la prolifération cellulaire de la lignée cellulaire CaCo2 par l'intermédiaire du récepteur au facteur de croissance épidermique, EGFR [81].

## **Complexité du système tumeur/immunité**

Comme nous l'avons vu, il existe une interaction entre le système immunitaire et les cellules tumorales. En effet, les tumeurs se développent dans un tissu composé de stroma et de cellules immunitaires. Le système immunitaire a des effets ambivalents sur la tumeur. L'inflammation promeut le développement tumoral tandis que les mécanismes d'immunosurveillance et d'immunoediting offrent une protection anti-tumorale. Ces effets, nombreux, sont le reflet de la complexité même du système immunitaire. Cette complexité s'illustre par la multiplicité des sous-populations qui infiltrent la tumeur. En réponse au système immunitaire, la tumeur module l'expression de gènes, en les induisant ou les inhibant. La complexité de la régulation de ces gènes est encore accrue avec la découverte récente des microARNs.

# Les microARNs

## Généralités

Depuis le début de la biologie moléculaire, il était acquis que l'expression d'un gène se faisait de façon linéaire : Le gène, sous le contrôle de différentes protéines régulatrices, semblait être tout d'abord transcrit en un ARN messager qui lui-même était traduit en protéines. Or, de petites séquences d'ARNs, appelée microARN, découverts récemment ont une grande importance dans le contrôle de ce processus. En effet, ils interviennent dans la régulation des ARN messagers et semblent être indispensables au bon fonctionnement des cellules. Ils interviennent dans de nombreux processus biologiques physiologiques (dont le fonctionnement des cellules du système immunitaire) et physiopathologiques (dont le cancer).

## Définition

Les microARNs sont des ARNs, non codants, de petites tailles (20 à 22 nucléotides environ). Ils sont présents dans le génome de nombreux organismes multicellulaires (plantes comme animaux) mais aussi chez les virus. Ce sont, pour la plupart, des régulateurs des ARN au niveau post-transcriptionnel. Ils interviennent dans de nombreux processus biologiques comme la prolifération, la différenciation cellulaire, l'apoptose... Un même microARN peut avoir des fonctions différentes en fonction du type cellulaire dans lequel il est exprimé. Le profil d'expression des microARNs est donc tissu-dépendant et peut être modifié dans un contexte pathologique.

Ils agissent sur leurs cibles grâce à la complémentarité imparfaite (chez les animaux) ou parfaite (chez les plantes) de leurs séquences. En effet, le mécanisme d'action des microARN chez les plantes nécessite une complémentarité parfaite alors que chez les animaux, d'autres mécanismes complémentaires autorisent une plus grande tolérance. Cette souplesse permet au miR de se fixer sur plusieurs cibles : Un microARN peut, en effet, se fixer sur des centaines d'ARNm. À l'inverse, une protéine peut être la cible de plusieurs microARNs.

## Historique

En 1993, l'équipe de V. Ambros en étudiant des gènes contrôlant le développement de la larve de *C. elegans* met en évidence l'existence de deux petits ARN *lin-4* et *let-7* [82]. Ces deux petits ARN sont décrits comme intervenant sur la régulation d'autres ARN. Cependant l'hypothèse envisagée est une action indirecte de ces petits ARN [83]. Ce n'est qu'en 2001 que le fonctionnement des

microARN (avec une action directe des petits ARN sur les ARNm) est décrit [84] [85]. En 2003 apparaît la première description du changement de profil des microARNs dans les cancers [86]. Depuis les années 2000, les microARNs ont été décrits chez les organismes des plus simples au plus complexes : Ils ont été étudiés chez les plantes, les algues, les champignons, les virus, les poissons, les mammifères... Cette propriété de conservation dans l'évolution a notamment servi pour l'identification des microARNs. Actuellement, on en dénombre plus de 800 décrits chez l'Homme mais plus de 1000 ont été prédits. Leur panel de fonctions s'agrandit encore tandis que leurs mécanismes d'action demeurent mal connus. Leurs cibles sont souvent putatives.

## **Conservation dans l'évolution**

Beaucoup de microARNs sont conservés dans l'évolution, ce qui suggère leurs rôles cruciaux et la pression de sélection qu'ils subissent. En effet, let-7 décrit pour son implication dans le développement chez *C. elegans*, est présent également chez l'Homme, la souris, la drosophile, le poisson-zèbre... [87] Les séquences-cibles des microARNs sont également conservées. Il est décrit qu'un changement (mutation, amplification, translocation, délétion...) du microARN ou de sa séquence-cible dans l'ARNm peut être la cause d'un cancer [88].

## **Les microARNs dans le génome**

La plupart des microARNs proviennent de régions du génome distinctes de celles où se trouvent des gènes déjà décrits. Cependant un quart des microARNs humains se trouvent dans les introns de gènes d'ARNm [89]. Il existe de rares exemples où le microARN se trouve dans un exon. Les microARNs ont, la plupart du temps, la même orientation que l'ARNm suggérant qu'ils sont transcrits de concert avec l'ARNm et n'ont donc pas de promoteur propre. Cela peut aboutir à l'expression coordonnée du microARN et de la protéine. Toutefois, un tiers des microARNs se trouvant dans les introns ont leurs régions d'initiation de la transcription propres et indépendants [90].

La plupart des microARNs humains sont isolés dans le génome, mais il existe des transcrits primaires multi-cistroniques appelés "cluster" [89] qui peuvent être co-exprimés ou être exprimés de manière indépendante. A l'inverse, des microARNs, co-exprimés, peuvent ne pas appartenir au même cluster.

## **Mécanismes de fonctionnement des microARNs**

### **La transcription et la maturation des microARNs**

Les microARNs sont transcrits sous la forme d'un microARN primaire ou pri-miARN. Ils sont généralement transcrits par l'ARN polymérase II et polyadénylés (Figure 6). L'ARN polymérase III peut également les transcrire mais, dans ce cas, ils ne seront pas polyadénylés. Un pri-miARN contient au moins une structure tige-boucle. Celle-ci est reconnue par le complexe enzymatique composé de la RNase III Drosha et de ses co-facteurs DGCR8. Ce processus conduit la maturation du pri-miARN en microARN précurseur ou pre-miARN. Ce dernier, plus court (une centaine de nucléotides), a égale-

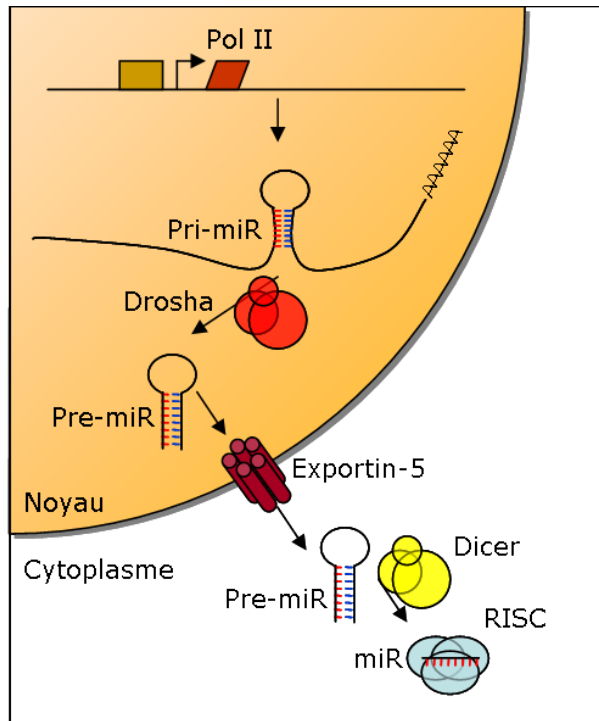


FIGURE 6 – La transcription et la maturation des microARNs

ment une structure en tige-boucle. Il est reconnu par l'exportine-5 puis transporté dans le cytoplasme. Il est pris en charge par une autre RNase III, Dicer, qui coupe la boucle. Un ARN de petite taille est alors libéré : Il contient le microARN sens (miR) et le microARN anti-sens (miR\*). Les deux brins sont probablement séparés sous l'action d'une hélicase encore inconnue à ce jour. Généralement, le microARN anti-sens est dégradé. Cependant, il arrive que les deux brins soient fonctionnels et pris en charge séparément. Le microARN mature, simple brin, est aussitôt pris en charge par un complexe nommé RISC contenant des protéines de la famille des Argonautes. C'est le complexe miRNA/RISC qui va agir sur les ARN messagers (Figure 6). Dans la plupart des cas, seul le microARN sens est pris en charge par le complexe RISC. Cependant il a été montré des cas où les deux brins sont pris en charge séparément et sont tous les deux fonctionnels.

### L'action des microARNs sur les ARN messagers

Les microARNs ciblent généralement la partie 3' des ARNs messagers avec une complémentarité imparfaite. Toutefois, un microARN se fixant sur la partie 5' d'un ARNm peut être tout aussi efficace. Les microARNs ont plusieurs types d'action sur les ARNm. En effet, ils peuvent activer ou inhiber la traduction. L'inhibition de la traduction peut se faire via une "simple" répression ou via la dégradation de l'ARN messenger.



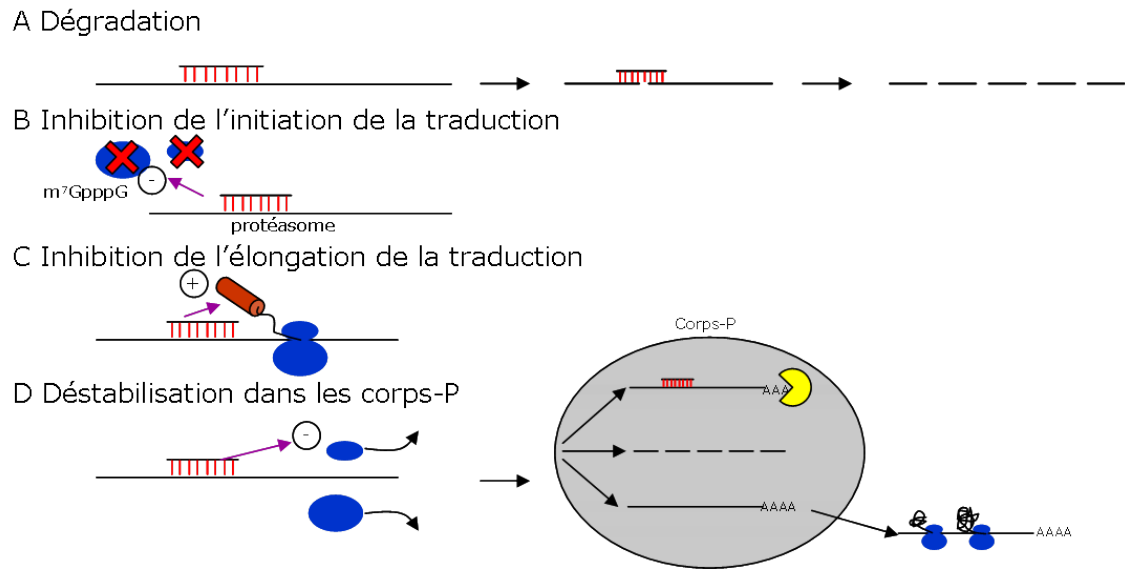


FIGURE 7 – A : Parfaite complémentarité entraînant une dégradation des ARNm ; B : Inhibition de l'initiation de la traduction ; C : Dégradation des polypeptides ; D : Inhibition de l'élongation

### Inhibition de la traduction

L'inhibition de la traduction a souvent lieu lorsque la complémentarité microARN/ARNm n'est pas parfaite. Elle peut avoir lieu à l'initiation et/ou au stade de l'élongation de la traduction. Trois différents mécanismes d'inhibition ont été décrits.

Le premier empêche l'initiation de la traduction. Elle nécessite que l'ARNm ait une structure 7-méthylguanoside en 5' (m<sup>7</sup>G cap). Le microARN interfère dans la reconnaissance de ce site par les protéines qui s'y lient (Figure 7B).

Le deuxième mécanisme intervient pendant le stade d'élongation, selon un mode encore peu connu. Une fraction au moins des ribosomes est engagée dans la traduction mais ne sont pas stables. Le modèle proposé est le suivant : Les ribosomes commencent la traduction mais se détachent prématurément [91]. Un autre modèle propose que les polypeptides naissants soient dégradés au fur et à mesure [92]. En effet, le complexe microARN/RISC semble être à même de recruter des enzymes protéolytiques qui dégradent les neopeptides au fur et à mesure de leurs productions (Figure 7C).

Le troisième mécanisme d'inhibition de la traduction pourrait être la déstabilisation de l'ARNm. Ce mécanisme serait lié à leurs capacités à recruter des enzymes enlevant la queue polyA des ARNm. Le complexe microARN/ARNm chapeauté par RISC, est envoyé dans une structure du cytoplasme appelé corps-P. Dans ces structures, on trouve des ARNm ciblés par des microARNs, des enzymes permettant d'enlever la queue polyA et les coiffes, et des RNAses. Toutefois, dans certaines circonstances comme un stress cellulaire, un ARNm peut ressortir du corps-P pour être à nouveau traduit. Le corps-P ne servirait alors que d'espace de stockage d'ARNm inactifs (Figure 7D) [92].

## Dégradation de l'ARN messager

Lorsque la complémentarité microARN/ARNm est parfaite ou quasi-parfaite, le clivage de l'ARNm est catalysé par la ribonucléase de RISC (Ago2 chez l'Homme) au centre de la région reconnue par le microARN (Figure 7A). Les fragments d'ARNm alors libérée sont ensuite dégradés. Chez les animaux, seul le miR-196 a été décrit comme utilisant cette voie. Il induit la dégradation de l'ARNm de Hoxb8 chez la souris [93]. En revanche, chez les plantes, cette voie est majoritaire.

## Autres rôles des microARNs

Il a été également montré que, dans certaines conditions, les microARNs peuvent activer la traduction de leurs ARNm cibles [94] ou interagir avec la transcription de gène [95] [96]. En effet, Vasudevan *et al.* ont montré que miR-369-3 s'associe avec les protéines AGO (Argonaute) et FXR1 (fragile-X-mental-retardation-related protein 1) pour se fixer au niveau de la région ARE (AU-rich element) du gène TNF $\alpha$  (tumor necrosis factor- $\alpha$ ) afin d'en activer sa traduction. D'autre part, Kim *et al.* ont découvert un microARN, miR-320, codé dans la région promotrice du gène POLR3D. Mir-320 interagit avec les protéines AGO1, PcG (polycomb group) et H3K27me3 (tri-methyl histone H3 lysine 27) au niveau du promoteur de POLR3D pour inhiber sa transcription. Un autre microARN, miR-373 cible aussi les promoteurs des gènes codant l'E-cadherine et CSDC2 (cold-shock domain-containing protein C2). Mais à la différence de miR-320, miR-373 active la transcription de ces deux gènes [96]. À ce jour, seuls ces trois exemples ont été décrits et les mécanismes précis ne sont pas connus.

## La régulation de l'expression des microARNs

La plupart des microARNs proviennent de régions du génome exemptes de tout gène (codant pour des longs ARNm) connu. Cependant un quart des microARNs sont liés à un gène codant pour un ARNm long : Généralement ils se trouvent dans les introns de ceux-ci. Un tiers d'entre eux, ont leurs régions d'initiation de la transcription propres et indépendantes [90]. Les deux tiers restant sont probablement exprimés en même temps que l'ARN dans lequel ils se trouvent. La plupart des microARNs ont donc des promoteurs indépendants et peuvent être régulés par des facteurs de transcription. Par exemple, c-Myc, un facteur de transcription souvent hyperactif dans les cellules cancéreuses, transactive directement le cluster miR-17-92 [97].

Hormis la régulation de la transcription, l'expression des microARNs est aussi régulé par d'autres mécanismes post-transcriptionnels. Des régulations ont été observées à tous les stades de maturation et d'action des microARNs.

Une régulation possible du microARN est leur dégradation. Au stade pri-microARN, certains peuvent être modifiés par des protéines de la famille des ADAR (adenosine deaminases acting on RNA). Cette modification entraîne un changement de stabilité du microARN ou une altération de la spécificité du microARN pour sa cible. Il est possible qu'un changement de cible soit induit [98] [99] [100].

Toujours au niveau du noyau, la maturation du pri-microARN en pre-microARN de certains microARNs (comme let-7 dans les cellules ES) peut être inhibé [92]. Bien que le mécanisme soit encore

inconnu, Drosha n'est plus capable de couper le pri-miRNA en pre-miRNA. Il existe plusieurs hypothèses qui ne s'excluent pas : Une protéine ou un autre microARN peut bloquer la reconnaissance du pri-miRNA par Drosha. Il est possible aussi que ce soit Drosha lui-même qui soit régulé.

Une fois que les micro-ARN sont matures, ils doivent être pris en charge par le complexe RISC. Or les protéines entrant dans la composition du complexe RISC peuvent elles aussi être régulées, et empêcher l'interaction RISC/miRNA. Ainsi chez la drosophile, la stimulation de neurones olfactifs entraîne la dégradation de la protéine Armitage. Chez la drosophile, Armitage est une protéine composante du complexe RISC. Sa dégradation entraîne la libération de l'ARNm du microARN/RISC levant ainsi son inhibition [101]. L'équivalent d'Armitage dans les cellules humaine, la protéine Dnd-1 (dead end 1), joue le même rôle [102].

Une fois associés avec le complexe RISC, les microARNs doivent reconnaître la séquence-cible sur les ARNm. Des protéines peuvent empêcher la fixation d'un microARN en se fixant sur l'ARNm : Par exemple, dans des conditions de stress, la protéine Hur (AU-rich element-binding protein) s'associe à la partie 3' de l'ARNm de CAT-1 empêchant ainsi la fixation du miR-122 dans les cellules de carcinome hépatique [103].

Un des mécanismes d'action du microARN est de transporter les ARNm dans les corps-P où la queue polyA est enlevée. Dans certaines conditions comme un stress cellulaire, les ARNm déjà lié avec un microARN peuvent sortir des corps-P et être à nouveau traduits.

## **Les microARNs et le système immunitaire**

Les microARNs participent à dans de nombreux processus biologiques. La différenciation des cellules du système immunitaire et leurs modes d'action impliquent des microARNs, tant dans le système immunitaire inné, qu'adaptatif.

### **Dans le système immunitaire inné**

#### **La différenciation du système immunitaire inné**

Les profils d'expression des microARNs dans les différentes cellules du système immunitaire inné sont assez peu étudiés. Toutefois les microARNs miR-223, miR-155, miR-106a, miR-17-92, miR-146 et miR-142 ont leurs expressions qui varient dans ces cellules.

Le miR-223 est spécifique de la lignée granulocyte (Figure 8). Son expression augmente durant la maturation de ces cellules sauf pour les mastocytes dans lesquels l'expression diminue. Il semble avoir à la fois un rôle de régulateur positif et négatif. Concernant son rôle positif, sa cible serait NFI-A (nuclear factor I-A). C/EBP $\alpha$  (une protéine CCAAT enhancer) et PU.I (un facteur de transcription spécifique de la lignée myéloïde) augmenterait sa transcription [104]. Concernant son rôle négatif, sa cible serait Mef2c, un facteur de transcription promouvant la différenciation et l'activation des granulocytes [105] [106].

L'expression du cluster miR-17-92 ainsi que de miR-106a est diminuée dans la différenciation des monocytes. Le cluster miR-17-92 code pour six microARNs différents : miR-17, miR-18a, miR-19a,

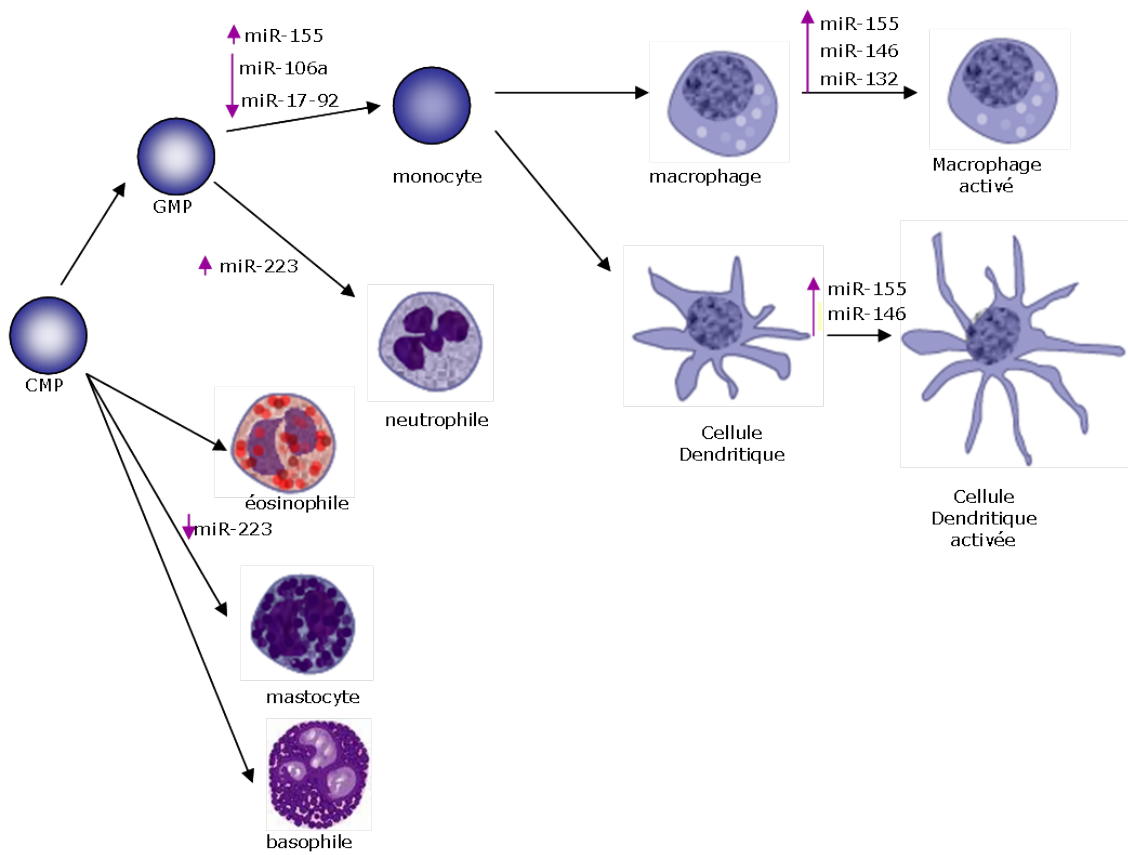


FIGURE 8 – Variation d'expression des microARNs dans les cellules du système immunitaire inné

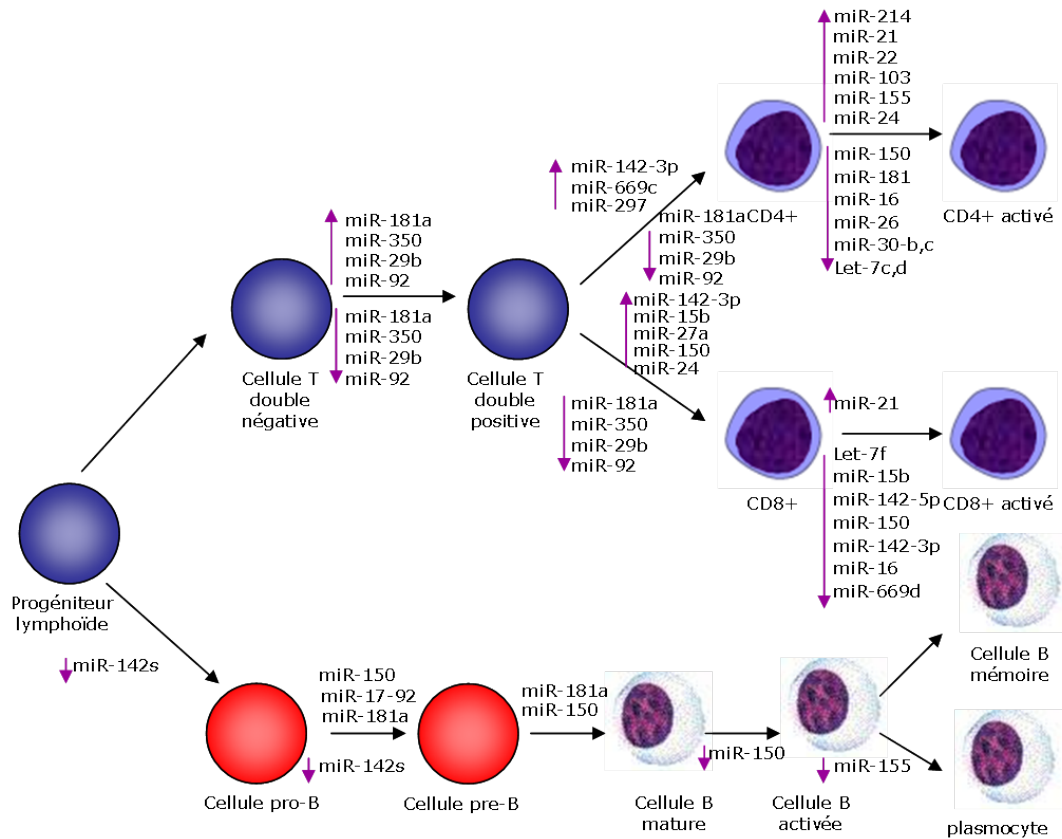


FIGURE 9 – Variation d’expression des microARNs dans les cellules du système immunitaire adaptatif. Adapté de Sonkoly *et al.* [108]

miR-20a, miR-19b et miR-92. Leurs rôles respectifs ne sont pas connus mais ils interviennent dans le contrôle de la prolifération et l’apoptose. Leur expression varie dans les cellules tumorales.

MiR-155 intervient dans beaucoup de cellules immunitaires. Cependant son rôle dans la différenciation des monocytes n’est pas encore connu.

### L’activation du système immunitaire inné

Le rôle des microARNs dans l’activation du système immunitaire inné est peu décrit. Seules les conséquences d’une activation des TLRs (Toll like receptor) ont été étudiées : L’activation des TLR2, TLR4 et TLR5 régule positivement miR-146, miR-132 et miR-155. miR-146a est induit par NF- $\kappa$ B et régule négativement IRAK1 (IL-1 receptor-associated kinase 1) et TRAF6 (TNF receptor-associated factor 6). miR-155 est également induit par l’activation des TLR3 et TLR9 mais aussi par la cytokine IFN- $\beta$ . Il induit une augmentation de la production du TNF- $\alpha$  et régule négativement FADD (Fas-associated death domain) entre autre [107].

## Dans le système immunitaire adaptatif

### Les lymphocytes B

Jusqu'à récemment, peu de microARN ont été décrits dans les différents stades de différenciation des lymphocytes B. Seul miR-150 est assez bien décrit (Figure 9). miR-150 a pour cible c-myc dont l'expression diminue durant la maturation des cellules B, parfait miroir de l'expression de miR-150. Basso *et al.* ont étudié récemment le miRNome des cellules B humaines. Les microARNs de la famille de miR-199, miR-125 et miR-138 ont des expressions distinctes dans les cellules B du centre germinatif [109]. Cependant leurs variations d'expression dans les différents stades de différenciation n'ont pas encore été publiées précisément. miR-155 est très important dans les cellules B et les plasmocytes : Il a un effet direct ou indirect sur la production des cytokines (notamment IL-2 et IFN- $\gamma$ ), la production d'IgM et d'IgG1. *In vitro*, miR-155 cible AID mais ce n'est pas sa seule cible dans les cellules B. Il agit également sur MAF [110] et PU.1 [111].

### Les lymphocytes T

Contrairement aux lymphocytes B, les variations d'expression des microARNs dans les différentes étapes de la différenciation des lymphocytes T sont assez bien décrites, même si les cibles des microARNs ne sont pas connues. Le miR-181a intervient dans les voies de différenciation précoce des lymphocytes T et B, sans que l'on connaisse son rôle exact dans ces étapes. Mais sa présence empêche le passage des lymphocytes T au stade double positive (CD4+ CD8+) au stade simple positive (CD4+ ou CD8+) car il inhibe des gènes nécessaires à la maturation comme Bcl-2, CD69 ou le TCR. Certains microARNs sont spécifiques d'un stade de différenciation précis. Les miR-350 et miR-92 sont fortement exprimés dans les cellules doubles positives contrairement aux autres stades. miR-669c et miR-297 sont spécifiques des cellules CD4+ et miR-15b, miR-150, miR-24 et miR-27a, des CD8+ [108]. Les microARNs sont également importants pour l'orientation de la réponse immunitaire. Par exemple, l'augmentation du microARN miR-155 et la légère augmentation du miR-146 dirigeront une réponse Th1 alors que l'augmentation des microARNs miR-233 et miR-146 pour les T régulateurs. Au contraire les lymphocytes Th2 présentent une très faible expression du miR-146 [106].

L'activation des lymphocytes T induit un changement d'expression de certains microARNs : Après leurs activations par CD3+ *in vitro*, les microARNs miR-21, miR-22, miR-24, miR-103, miR-155, miR-204 sont sur-exprimés tandis que d'autres (miR-16, miR-26, miR-30b,c, miR-150, miR-181 et let-7) sont sous-exprimés [108].

## MicroARN et cancer

Dès le début de leur découverte, la possible implication des microARNs dans le cancer était soutenu par certaines observations. Tout d'abord, les microARNs sont impliqués dans le contrôle de la prolifération et de l'apoptose des cellules, deux mécanismes cruciaux impliqués dans le cancer. De plus, beaucoup de gènes codant pour les microARNs se trouvent dans des régions du génome connus pour être fréquemment amplifiées ou délétées dans les cancers humains. Enfin le profil d'expression

	Oncogène	gène suppresseur de tumeur
oncomiR	+	-
microARN suppresseur de tumeur	-	+

FIGURE 10 – Effets possibles des microARNs sur les oncogènes et gènes suppresseurs de tumeur

des microARNs dans les tissus sains et dans les tissus tumoraux est différent. Deux grandes catégories de microARN ont été définies par rapport à leurs effets dans le cancer : Les microARN pro tumoraux appelés oncomiRs et les microARNs suppresseurs de tumeurs.

## Les oncomiRs

### Définition

Les oncomiRs sont des microARNs qui agissent comme des oncogènes. Un oncomiR peut agir de plusieurs manières : Son expression peut aboutir soit à activer un oncogène, soit à inhiber un gène suppresseur de tumeur (Figure 10). De nombreux microARNs ont été décrits comme oncomiR dans les différents types de cancer (Tableau 2).

### miR-17-92

Le cluster de microARN miR-17-92 est le premier cluster décrit comme oncomiR. Il comprend six microARNs : miR-17, miR18a, miR-19a, miR19b-1, miR-20a et miR-92-1. Il est localisé sur le chromosome 13, dans une région où des amplifications sont décrites comme étant impliquées dans des cas de lymphomes et de tumeurs solides. Ce cluster de microARN a d'abord été décrit comme intervenant dans la régulation du cycle cellulaire, puis dans l'induction de l'angiogenèse [117].

Le facteur de transcription c-Myc, souvent hyperactif dans les cellules cancéreuses, transactive directement le cluster miR-17-92 [97]. E2F1, E2F2 et E2F3 qui induisent l'expression de gènes permettant la transition des phases G1-S du cycle cellulaire, ont été les premières cibles vérifiées du cluster [97][118][119]. De plus, E2F1 et E2F3 peuvent directement activer la transcription du cluster miR-17-92, formant ainsi une boucle de régulation négative. Or de forts taux de ces protéines peuvent induire l'apoptose. La boucle de régulation négative permet de réguler les protéines E2F promouvant ainsi la division cellulaire plutôt que la mort cellulaire.

Un autre régulateur du passage G1-S du cycle cellulaire, CDKN1A (cyclin dépendant kinase inhibitor) ou p21 est régulés par miR-17, miR-20a, miR106b et miR-93 [120]. La sur-expression de ces microARNs suffit à outrepasser l'arrêt du cycle cellulaire induit par des dommages de l'ADN. Bim, une protéine proapoptotique qui antagonise des protéines antiapoptotiques comme Bcl-2, est aussi une cible des miR-17-92 [121].

Toutes ses données montrent l'activité pro-tumorale de ces microARNs. Cependant, la perte de fonction de ce cluster peut présenter un avantage pour la tumeur dans certains cas. En effet, une perte de fonction a été observée dans 16% des cancers de l'ovaire, 22% des cancers du sein et 20% des

microARN	Type de cancer/ localisation	Changement du niveau d'expression dans les cellules cancéreuse	Cible	Références
miR-9	cerveau	diminué	granuphilin Slp4	[112]
miR-10b	lignée de cancer du sein	augmenté	HOXD10	[112]
miR-17-92	lymphome, colon, poumon, sein, pancréas, prostate, estomac	augmenté	E2F1, p21, BIM, Tsp1, CTGF	[113] [112] [114]
miR-18	carcinome hépatique	augmenté	CTGF, RANKL	[114]
miR-21	glyoblastome, sein, pancréas, colon, foie	augmenté	BCL-2, PTEN, TPM1, PDCD4, RECK	[115] [113] [112] [114][116]
miR-34a	leucémie	augmenté	?	[112]
miR-133	(?)	augmenté	SRF, Caspase 9	[112]
miR-142	lymphome B	diminué	?	[113][112]
miR-146b	carcinome, thyroïde	augmenté	IRAK1, TRAF6, Kit	[113][112]
miR-155	lymphomes, tumeurs multiples	augmenté	AGTR1, TP53INP1	[113] [113]
miR-181a	lymphome	augmenté	TCL1, OXA11, SHP2, DUSP6, DUSP5, PTPN-22	[112]
miR-195	lymphomes	augmenté	NGN3	[112]
miR-221	glioblastome, thyroïde, thyroïde, leucémie	augmenté augmenté	Kit	[113]
miR-222	leucémie, thyroïde, foie	augmenté	Kit	[113][116]
miR-223	vessie	augmenté	(?)	[112]
miR-372	testicule	augmenté	LATS2	[113] [112] [114]
miR-373	testicule	augmenté	LATS2	[113] [112] [114]
miR-106-363	leucémie, tumeurs multiples	augmenté	AML1, M-CSFR, Rb, Mylip, Rbp1-like, HIPK3	[113]

TABLE 2 – Exemples d'oncomiRs et leurs cibles



mélanomes. Dans les lignées de cancer du sein, miR-17 réduit la prolifération des cellules. Cet effet est dû en partie à l'inhibition du gène AIB1 (amplified in breast cancer 1) qui code un co-activateur transcriptionnel de E2F1 [122].

L'effet pro ou anti-tumoral de ces microARNs dépend donc de l'expression de leurs cibles dans une situation précise.

Un autre rôle des microARNs est dû à c-Myc qui a un potentiel inducteur d'angiogenèse au niveau de la tumeur. Son activation induit une diminution de protéines antiangiogéniques comme Tsp1 (thrombospondin-1) ou CTGF (connective tissue growth factor). Cette action est au moins due en partie au cluster de microARN : Tsp1 et CTGF sont tous les deux régulés par miR-17-92 [123]. Toutefois un autre microARN du cluster, miR-92a, cible des ARNm pro-angiogéniques, dont la sous-unité  $\alpha 5$  de l'intégrine [124]. Cela montre encore une fois que tous les microARNs de ce cluster n'ont pas des rôles similaires.

### **Les oncomiRs et les cancers colorectaux**

Dans les cancers colorectaux, les microARNs miR-20, miR-21, miR-29a, miR-29b, miR-17-5p, miR-19a, miR-15b, miR-30c, miR-92, miR-106a, miR-107, miR-148, miR-181b, miR-191, miR-200b, miR-200c, miR-221 et miR-223 sont sur-exprimés [125].

Le microARN miR-21 a pour cible PDCD4 (programmed cell death gene 4). Cela a pour conséquence de promouvoir la migration cellulaire, l'invasion et les métastases [126].

De nombreux microARNs sont également montrés comme étant prédictifs de la survie. Par exemple, miR-20a, miR-21, miR-106a, miR-181b ou miR-203 sont associés à un mauvais pronostic. miR-106a corrèle également avec le stade de la tumeur et la survie globale [127].

Il est intéressant de noter que la famille des miR-200 (miR-200a, miR-200b, miR-200c, miR-141 et miR-429) ciblent ZEB1 et ZEB2. ZEB1 a été montré comme régulant la différenciation T et l'expression de CD4 et réprimant la production de la cytokine IL-2 [128].

Un autre microARN semble être lié à la réponse immunitaire anti-tumorale. En effet, miR-29 cible CD276. CD276 est préférentiellement exprimé par les tissus tumoraux. Une de ses formes (la forme 2Ig) promeut l'activation des lymphocytes T et la production IFN- $\gamma$  en étant reconnu par un récepteur, inconnu à ce jour, présent sur les lymphocytes T activés. Le miR-29 empêche alors cette activation des lymphocytes T en inhibant l'expression de CD276 [129].

### **Les microARNs suppresseurs de tumeurs**

#### **Définition**

Les microARNs suppresseurs de tumeur peuvent agir de plusieurs manières : Leur expression peut aboutir soit à inhiber un oncogène, soit à activer un gène suppresseur de tumeur (Figure 10). De nombreux microARNs agissent comme des suppresseurs de tumeur dans les différents types de cancer (Tableau 3).

microARN	Type de cancer/ localisation	Changement du niveau d'expression dans les cellules cancéreuse	Cible	Références
famille let-7	poumon, colon	diminué	Ras, PRDMI, HMGA2, IGF2BP1, c-Myc, NF2, LIN28b	[113] [112]
miR-10b	sein	diminué	?	[113]
miR-15a	lymphome, adénome, poumon	déleté	BCL-2, NGN3, Rb	[113][112] [130]
miR-16-1	lymphome, adénome	déleté	BCL-2, NGN3	[113][112]
miR-17-5p	sein	diminué	AIB1, E2F1	[113]
miR-17-92	sein	diminué	AIB1	[112]
miR-20a	?	diminué	E2F1	[113]
miR-29b	leucémie, poumon	diminué	MCL-1, TCL-1	[113][112]
miR-31	sein	diminué	RhoA	[131]
miR-34a	colon, poumon, neuroblastome	diminué	CDK4, E2F3, BCL-2	[112] [114]
miR-34b/c	poumon	diminué	CDK6, c-Met, c-Myc	[112]
miR-122	carcinome hépatique	diminué	CCNG1, CAT-1b	[112]
miR-124a	colon, sein, poumon, leucémie, lymphome	diminué	CDK6	[113]
miR-125b	sein	diminué	?	[113]
miR-126	sein	diminué	VCAM-1	[112]
miR-127	vessie, lymphome	diminué	BCL-6	[113]
miR-143	colon, leucémie	diminué	KRAS	[132] [113][112]
miR-145	colon, sein	diminué	?	[113]
miR-181b	leucémie	diminué	Tcl-1	[113]
miR-335	sein	diminué	SOX4, Tenascin C Jagged-1	[112]

TABLE 3 – Exemples de microARNs suppresseurs de tumeur et leurs cibles

## let-7

Le premier membre de la famille des let-7 a été décrit chez *C. elegans*. Il induit la sortie du cycle cellulaire et la différenciation terminale dans un type particulier de cellules durant la transition larve/adulte. Chez l'Homme, il existe une douzaine de microARNs de la famille de let-7 dans 8 loci différents. Les cibles prédites de ces douze microARNs sont similaires, pourtant leurs fonctions ne sont pas identiques. Les cibles physiologiques de cette famille sont des gènes embryonnaires. Normalement, cette famille est fortement exprimée à un stade tardif du développement embryonnaire. Dans les cancers, ils sont exprimés dans des cellules à un stade précoce de dé-différenciation. L'hypothèse est que la famille let-7 contrôle des gènes oncofoetaux afin de prévenir le processus de dé-différenciation. Beaucoup de gènes de la famille let-7 sont situés dans des zones de chromosome délétés dans de nombreux cas de cancers. D'une manière générale, le niveau d'expression de cette famille de microARN est diminué dans les cancers.

Les rôles des différents membres de la famille let-7 sont difficiles à distinguer les uns des autres. Les différents gènes codant pour let-7 sont directement réprimés par c-Myc. L'ARN précurseur des microARNs peut également être régulé par Lin-28b [133] qui empêche Drosha de transformer le pri-miR en pre-miR. let-7 a pour cible Ras [134], HMGA2 (high mobility group A2) [135] et c-Myc lui-même [136]. Il est intéressant de noter que HMGA2 est un gène embryonnaire précoce non exprimé dans la plupart des tissus adultes. En revanche, il est surexprimé chez l'adulte dans de nombreux cancers comme le cancer pancréatique, du poumon, le carcinome squameux ou le neuroblastome. Tous les membres de la famille de let-7 n'ont pas la même fonction. Let-7a peut avoir des fonctions oncogéniques. En effet, la cytokine pro-inflammatoire IL-6 conduit à l'expression de let-7a, ce qui aboutit à une croissance de la tumeur et à la résistance aux drogues [137].

## Les microARNs suppresseurs de tumeur et les cancers colorectaux

Dans les cancers colorectaux, les microARNs miR-34a, miR-126, miR-143, miR-145 et miR-342 sont souvent décrits comme étant sous-exprimés dans la tumeur. L'expression de certains microARNs semble lié à un pronostic : miR-145 et miR-320 sont décrits comme étant de bon pronostic [138]. Une cible putative de miR-145 est IRS-1 (Insulin Receptor substrate 1). IRS-1 transduit des signaux mitogéniques, anti-apoptotiques et de dé-différenciation [139]. Son inhibition par miR-145 correspond alors à un processus anti-tumoral.

# **Objectifs**

Aujourd'hui les cancers colorectaux sont classés suivant des critères anatomopathologiques : l'extension pariétale de la tumeur (stade T), l'existence et le nombre de ganglions lymphatiques envahis (stade N) et la présence ou non de métastases (stade M). Grâce à ces trois paramètres, les patients sont classés en quatre groupes. Cette classification a pour but de regrouper les différents stades de pronostic assez proches. Cependant elle est imparfaite car les groupes de patients sont assez peu homogènes. Il est donc important de trouver d'autres marqueurs pronostiques plus performants.

Le système immunitaire a un rôle important dans le contrôle de la progression tumorale : Divers types de cellules immunitaires ont été mis en évidence au site de la tumeur et la présence de lymphocytes infiltrant la tumeur joue un rôle dans l'inhibition de la croissance tumorale. Cependant la présence d'une tumeur malgré la capacité du système immunitaire à la détruire, met en évidence des mécanismes d'échappement.

Mon objectif est d'étudier les différentes populations immunitaires et les mécanismes d'échappements tumoraux lors de la progression de la tumeur à travers les tissus (stade T). Notre laboratoire a identifié la population lymphocytaire T mémoire comme étant associée à une absence de signes précoces d'invasion métastatique et à une survie améliorée [19]. Nous avons poursuivi ce travail en analysant différentes populations immunitaires afin de déterminer leurs valeurs pronostiques respectives et leurs coordinations.

Afin de compléter les études des ARNm et protéines, je me suis intéressée aux microARNs. Les microARNs sont impliqués dans différents processus cellulaires comme la prolifération, l'apoptose ou le fonctionnement des cellules immunitaires, ainsi que dans des pathologies comme le cancer. J'ai comparé le profil d'expression des microARNs dans les tissus tumoraux et dans les tissus adjacents sains. J'ai évalué la valeur pronostique de chaque microARN et établi des corrélations avec les différentes populations immunitaires.

# Résultats

## **Article 1**

# **"Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome"**

Jérôme Galon, Anne Costes, Fatima Sanchez-Cabo, Amos Kirilovsky, Bernhard Mlecnik, Christine Lagorce-Pagès, Marie Tosolini, Matthieu Camus, Anne Berger, Philippe Wind, Franck Zinzindohoué, Patrick Bruneval, Paul-Henri Cugnenc, Zlatko Trajanoski, Wolf-Herman Fridman, Franck Pagès

Science, 2006, volume 313, pages 1960-64

Notre laboratoire a identifié la population lymphocytaire T mémoire comme étant associée à une absence de signes précoces d'invasion métastatique et à une survie améliorée [19]. Nous avons approfondi ce travail en recherchant les paramètres immunologiques aux potentiels prédictifs les plus puissants en terme de survie des patients atteints de cancers colorectaux. Nous avons comparé l'importance de la réponse immunitaire par rapport au poids des critères d'invasion tumorale (T, N, M) sur le pronostic des patients.

J'ai effectué le marquage des lymphocytes T CD3 par immunohistochimie sur les Tissue MicroArrays de la cohorte de validation (figure 2F). J'ai analysé ce marquage à l'aide du logiciel Spotbrowser dans les deux zones tumorales (centre et front de progression).

## Résumé

Les paramètres anatomopathologiques sur lequel se base classiquement le pronostic des patients atteints de cancer colorectaux ne permettent pas une identification suffisamment précise des patients à risque. Par conséquent, certains patients ne bénéficient pas d'un traitement efficace. Or depuis quelques décennies, les recherches ont démontré que le système immunitaire jouait un rôle important dans le contrôle de la progression tumorale dans de nombreux cancers chez l'Homme et en particulier le cancer colorectal. Mais le rôle de la réponse immunitaire adaptative dans le contrôle de la progression tumorale et la récurrence est controversé.

Nous avons caractérisé les cellules immunitaires infiltrant la tumeur dans une large cohorte de cancers colorectaux humains. Nous avons analysé la réponse immunitaire adaptative cytotoxique de type Th1 en étudiant le niveau d'expression de gènes par RT-PCR quantitative et la densité des cellules immunitaires par immunohistochimie.

Ces données immunologiques, incluant le type, la densité et la localisation des cellules immunitaires, permettent de mieux prédire la survie des patients que les données histologiques utilisées actuellement.

L'analyse des corrélations des niveaux d'expression des gènes immunitaires montre un cluster dominant de gènes de type Th1 co-modulés (CD3 $\zeta$ , CD8 $\alpha$ , T-Bet, IRF1, IFN $\gamma$ , granulysine et granzyme B). L'expression coordonnée de ces gènes est inversement corrélée à la récurrence des patients.

Nous avons analysé par immunohistochimie en Tissue-MicroArray (TMA) dans deux zones tumorales (centre de la tumeur et front de progression) les marqueurs lymphocytaires CD3 et CD45RO. La coordination des densités de cellules CD3+ dans les deux zones tumorales analysées apporte une valeur pronostique puissante en terme de survie globale ( $p < 10^{-7}$ ) et de survie sans récurrence ( $p < 10^{-8}$ ) des patients bien supérieure à celle des paramètres anatomopathologiques (stade I, II et III). Ces observations ont été confirmées de façon indépendante sur deux autres cohortes de patients. L'analyse combinatoire des densités de cellules CD3+ et des cellules CD45RO+ a permis l'identification précise et sans ambiguïté des patients à risque élevé de récurrence ( $p < 10^{-4}$ ) indépendamment des paramètres anatomopathologiques.

Cette étude montre pour la première fois le caractère pronostic puissant des paramètres immunologiques (type, densité et localisation des cellules immunitaires) chez les patients atteints de cancers colorectaux. En tenant compte de ces paramètres, les critères d'invasion tumorale (TNM) ne sont plus significatifs et dépendent même des paramètres immunitaires.



mantle viscosity and crust thickness) in that region, especially over the Hudson Bay and Scandinavia, two prominent PGR active areas. It is possible that the ICE5G PGR model (13) may underestimate the PGR contribution to GRACE-observed ice mass loss over Greenland. However, the uncertainty of the estimated PGR contribution will not likely account for a significant portion of the  $-239 \pm 23$  km<sup>3</sup>/year ice mass loss observed by GRACE. If we adopt this ICE5G-based PGR contribution of mass rate over Greenland (about  $-5$  km<sup>3</sup>/year, with uncertainty at 100% of the signal, i.e.,  $\pm 5$  km<sup>3</sup>/year), then our GRACE estimate of Greenland ice mass rate is about  $-234 \pm 24$  km<sup>3</sup>/year.

The current GRACE estimate is significantly larger than an earlier estimate ( $-82 \pm 28$  km<sup>3</sup>/year), based on just the first 2 years of data (13). The difference is attributed both to increased melting in the most recent 1.5-year period and to improved filtering and estimation techniques (including use of numerical simulations), and the latter may have played a more important role. Increased recent melting may represent simple interannual variability or accelerated melting driven by steady Arctic warming (20). Despite close agreement between our GRACE estimate and recent radar interferometry estimates (2), quantification of Greenland ice mass balance remains a challenge. For example, another study (21) based on 10 years of radar altimetry data during the period 1992 to 2002 suggests a small mass gain for Greenland ( $\sim 11 \pm 3$  km<sup>3</sup>/year) (2), opposite in sign to the more recent estimate (2). On the other hand, thermomechanical ice models forced by general circulation model climate scenarios predict significant Greenland ice loss in the 21st century (22).

The numerical simulation approach used in this study is useful in interpreting GRACE time-variable gravity fields. It contrasts with the basin kernel function approach (13, 15), in which the focus is on a continent-wide average. Numerical simulations are useful in quantifying spatial leakage of variance and in testing hypotheses concerning possible regional contributors to change, such as the Southeast Glacier or Svalbard. Many error sources may affect our GRACE estimates, which include the remaining GRACE measurement error (after spatial smoothing), uncertainty in the background geophysical models used in GRACE (e.g., the uncorrected ocean pole effect in the release-01 GRACE data and errors in the atmospheric and ocean models over Greenland and surrounding regions), and unquantified other leakage effects.

The conclusion that ice loss has accelerated in recent years is independent of uncertainty in PGR effects, because, regardless of magnitude, PGR should contribute a constant rate to time series of any length. GRACE clearly detects a rate change in the most recent period, suggesting a contribution of about 0.54 mm/year to global sea level rise, well above earlier assessments (23). Time series are still relatively short,

and an understanding of interannual variation in ice mass rates is lacking for Greenland. Without question, the extension of the GRACE mission beyond 2010, or the development of a follow-up mission, will contribute fundamentally to separating contributions of ice mass change from other geophysical signals (such as PGR) that contribute to the observations.

#### References and Notes

1. W. Krabill *et al.*, *Geophys. Res. Lett.* **31**, L24402 10.1029/2004GL021533 (2004).
2. E. Rignot, P. Kanagaratnam, *Science* **311**, 986 10.1126/science.1121381 (2006).
3. E. Rignot, D. Braaten, S. P. Gogineni, W. B. Krabill, J. R. McConnell, *Geophys. Res. Lett.* **31**, L10401 10.1029/2004GL019474 (2004).
4. B. D. Tapley, S. Bettadpur, M. M. Watkins, C. Reigber, *Geophys. Res. Lett.* **31**, L09607 10.1029/2004GL019920 (2004).
5. Ch. Reigber *et al.*, *J. Geodyn.* **39**, 1 (2005).
6. S. Bettadpur, *Level-2 Gravity Field Product User Handbook*, The GRACE Project (Jet Propulsion Laboratory, Pasadena, CA, 2003).
7. J. Wahr, S. Swenson, V. Zlotnicki, I. Velicogna, *Geophys. Res. Lett.* **31**, L11501 10.1029/2004GL019779 (2004).
8. B. D. Tapley, S. Bettadpur, J. Ries, P. F. Thompson, M. M. Watkins, *Science* **305**, 503 (2004).
9. R. Schmidt *et al.*, *Global Planet. Change* **50**, 112 (2006).
10. D. P. Chambers, J. Wahr, R. S. Nerem, *Geophys. Res. Lett.* **31**, L13310 10.1029/2004GL020461 (2004).
11. J. L. Chen, C. R. Wilson, B. D. Tapley, J. S. Famiglietti, M. Rodell, *J. Geodesy* **79**, 532 10.1007/s00190-005-9 (2005).
12. W. R. Peltier, *Annu. Rev. Earth Planet. Sci.* **32**, 111 10.1146/annurev.earth.32.082503.144359 (2004).
13. I. Velicogna, J. Wahr, *Geophys. Res. Lett.* **32**, L18505 10.1029/2005GL023955 (2005).
14. M. E. Tamisiea, E. W. Leuliette, J. L. Davis, J. X. Mitrovica, *Geophys. Res. Lett.* **32**, L20501 10.1029/2005GL023961 (2005).
15. I. Velicogna, J. Wahr, *Science* 10.1126/science.1123785 (2006).
16. J. L. Chen, B. D. Tapley, C. R. Wilson, *Earth Planet. Sci. Lett.* **248**, 353 (2006).
17. G. Ramillien *et al.*, *Global Planet. Change* **53** (no. 3), 198 (2006).
18. J. L. Chen, C. R. Wilson, K.-W. Seo, *J. Geophys. Res.* **111**, B6, B06408, 10.1029/2005JB004064 (2006).
19. T. Sato *et al.*, *Geophys. J. Inter.* **165**, 729 (2006).
20. F. S. Chapin *et al.*, *Science* **310**, 657 10.1126/science.1117368 (2005).
21. H. J. Zwally *et al.*, *J. Glaciol.* **51**, 509 (2005).
22. P. Huybrechts, J. Gregory, I. Janssens, M. Wilde, *Global Planet. Change* **42**, 83 (2004).
23. J. A. Church *et al.*, in *Climate Change: The Scientific Basis*. Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change, J. T. Houghton, Ed. (Cambridge Univ. Press, Cambridge, 2001), pp. 639–694.
24. The authors would like to thank the three anonymous reviewers for their insightful comments. This study was supported by the NASA Solid Earth and Natural Hazards and GRACE Program (under grants NNG04G060G, NNG04GP70G, and NNG04GF22G).

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/1129007/DC1  
SOM Text

Fig. S1

20 April 2006; accepted 7 July 2006

Published online 10 August 2006;

10.1126/science.1129007

Include this information when citing this paper.

## Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome

Jérôme Galon,<sup>1\*†</sup> Anne Costes,<sup>1</sup> Fatima Sanchez-Cabo,<sup>2</sup> Amos Kirilovsky,<sup>1</sup> Bernhard Mlecnik,<sup>2</sup> Christine Lagorce-Pagès,<sup>3</sup> Marie Tosolini,<sup>1</sup> Matthieu Camus,<sup>1</sup> Anne Berger,<sup>4</sup> Philippe Wind,<sup>4</sup> Franck Zinzindohoué,<sup>5</sup> Patrick Bruneval,<sup>6</sup> Paul-Henri Cugnenc,<sup>5</sup> Zlatko Trajanoski,<sup>2</sup> Wolf-Herman Fridman,<sup>1,7</sup> Franck Pagès<sup>1,7†</sup>

The role of the adaptive immune response in controlling the growth and recurrence of human tumors has been controversial. We characterized the tumor-infiltrating immune cells in large cohorts of human colorectal cancers by gene expression profiling and in situ immunohistochemical staining. Collectively, the immunological data (the type, density, and location of immune cells within the tumor samples) were found to be a better predictor of patient survival than the histopathological methods currently used to stage colorectal cancer. The results were validated in two additional patient populations. These data support the hypothesis that the adaptive immune response influences the behavior of human tumors. In situ analysis of tumor-infiltrating immune cells may therefore be a valuable prognostic tool in the treatment of colorectal cancer and possibly other malignancies.

**T**umors in mice and humans often contain infiltrates of immune cells. Experiments with immune-deficient mice have provided data supporting the role of adaptive immunity in cancer immunosurveillance (1–4). Tumor cells can express antigens and become targets for a T cell–mediated adaptive immune response (5, 6). The differentiation of naïve CD4<sup>+</sup> T cells

into T helper type 1 (T<sub>H</sub>1) cells producing interferon gamma (IFN- $\gamma$ ) promotes CD8 T cell–mediated adaptive immunity (7). In mice, immune cells appear to prevent the development of tumors and inhibit tumor progression (1, 3, 4). Anti-tumor immunity also leads to immunoediting, a process favoring the eventual outgrowth of tumor cells with reduced immunogenicity (3).

The role of immune cells in human neoplasia is less clear (8). Immune cells can release inflammatory mediators with proangiogenic and prometastatic effects (9–14). Tumor-infiltrating lymphocytes in melanoma (15), colorectal cancers (CRCs) (16–18), and ovarian cancers (19, 20) have been shown to inhibit tumor growth and are associated with improved prognoses. After antigen stimulation, a small population of antigen-specific memory T cells remains in the tissues (21). We recently showed that human CRCs with a high density of infiltrating memory and effector memory T cells were less likely to disseminate to lymphovascular and perineural structures and to regional lymph nodes (22). Using the same cohort of patients, we investigated the relationship between the type, density, and location of immune cells within tumors and the clinical outcome of the patients.

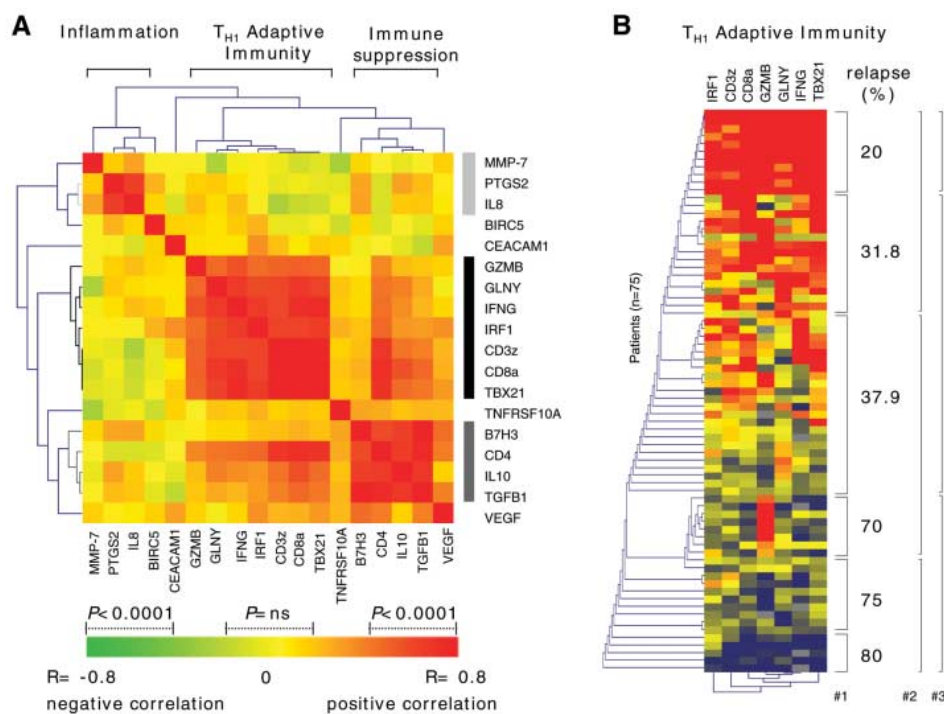
To this end, we conducted genomic and in situ immunostaining analyses on tumors from 75 and 415 patients, respectively (table S1). The data were entered into a dedicated Tumoral MicroEnvironment Database (TME.db; access available upon request). We used quantitative real-time polymerase chain reaction to evaluate the expression levels of genes related to inflammation, T<sub>H</sub>1 adaptive immunity, and immunosuppression. These genes showed variable expression patterns in the 75 tumors studied (fig. S1). Correlation analyses performed between all genes showed 39 highly significant combinations ( $P < 0.0001$ ) (fig. S1 and table S2). We identified a dominant cluster of co-modulated genes for T<sub>H</sub>1 adaptive immunity [genes encoding T-box transcription factor 21, interferon regulatory factor 1, IFN- $\gamma$ , CD3- $\zeta$ , CD8, granulysin, and granzyme B (GZMB)] (Fig. 1A). A hierarchical tree structure classifying the patients according to the expression levels of genes from this cluster revealed an inverse correlation between the expression of these genes and tumor recurrence ( $P$  value comparing patient groups, all  $P < 0.05$ ) (Fig. 1B). These data suggest that T<sub>H</sub>1 adaptive immunity has a beneficial effect on clinical outcome.

We next used tissue microarrays to investigate the in situ adaptive immune response in the center of the tumor (CT) and the invasive margin (IM) of 415 CRCs. Immunostainings for total T

lymphocytes (CD3), CD8 T cell effectors and their associated cytotoxic molecule (GZMB), and memory T cells (CD45RO) were quantified with the use of a dedicated image analysis workstation (Fig. 2A and figs. S2 to S4). Tumors from patients without recurrence had higher immune cell densities (CD3, CD8, GZMB, and CD45RO) within each tumor region (CT and IM), than did those from patients whose tumors had recurred (Fig. 2B). In each tumor region (CT and IM) and for each marker (CD3, CD8, GZMB, and CD45RO), there was a statistically significant correlation between immune cell density and patient outcome for a large range of cutoff values (fig. S5). In particular, using the cutoff that yielded the minimum  $P$  value for disease-free survival, the densities of CD3<sup>+</sup>, CD8<sup>+</sup>, GZMB<sup>+</sup>, and CD45RO<sup>+</sup> cells in each tumor region (CT and IM) allowed the stratification of patients into groups with different disease-free survival rates [ $P$  values corrected after (23), ranging from  $1.0 \times 10^{-2}$  to  $4.8 \times$

$10^{-6}$ ] and overall survival rates ( $P$  values ranging from  $5.5 \times 10^{-3}$  to  $7.9 \times 10^{-8}$ ) (Fig. 2C and tables S3 and S4). Reanalyses of the data using 100 repetitions of twofold cross-validations after (24) (tables S3 and S4) or setting the cutoff at the median of the data sets (tables S5 and S6) provided concordant results as to the prognostic value of each immune parameter.

We investigated whether the combined analysis of tumor regions could improve the prediction of patient survival. For all the markers of adaptive immunity (CD3, CD8, GZMB, and CD45RO), the combined analysis of CT plus IM regions [high density in both regions (HiHi) versus low density in both regions (LoLo)] increased the accuracy of prediction of disease-free and overall survival time for the different patient groups, as compared to single-region analysis (Hi versus Lo) (Fig. 2, D to F; figs. S6 and S7; and tables S3 to S6). Data were also analyzed using twofold cross-validation after



**Fig. 1.** (A) Correlation analyses performed between the 18 immunogenes were uploaded into the Genesis clustering program (28–30). A correlation matrix followed by unsupervised hierarchical clustering (Pearson uncentered algorithm) is represented from  $R = -0.8$  negative correlation (green) to  $R = 0.8$  positive correlation (red). For all correlations with  $0.4 < R < 0.9$ ,  $P < 0.05$  (table S1). The correlation matrix reveals a dominant cluster of co-modulated genes for T<sub>H</sub>1 adaptive immunity and two clusters of genes encoding mediators of inflammation and immunosuppression. (B) Hierarchical tree structure classifying the 75 patients according to the mRNA levels of the seven genes from the T<sub>H</sub>1 adaptive cluster, from maximal (red) to minimal (blue) expression levels. The percentage of patients with tumor recurrence (relapse) is indicated. Patients with a homogeneous increased expression of genes for T<sub>H</sub>1 adaptive immunity had the best prognosis. Log-rank tests comparing the disease-free survival times between patient groups reached statistical significance ( $P < 0.05$  for numbers 1, 2, and 3). In contrast, expression levels of inflammatory and immunosuppressive genes showed no correlation with tumor recurrence. MMP-7, matrix metalloproteinase 7; PTGS2, prostaglandin-endoperoxide synthase 2; IL8, interleukin-8; BIRC5, baculoviral IAP repeat-containing 5 (survivin); CEACAM1, carcinoembryonic antigen-related cell adhesion molecule 1; GLNY, granulysin; IRF1, interferon regulatory factor 1; TBX, T-box 21 (TBET); TNFRSF10A, tumor necrosis factor receptor superfamily, member 10a (TRAILR1); TGF $\beta$ 1, transforming growth factor- $\beta$  1; VEGF, vascular endothelial growth factor.

<sup>1</sup>INSERM U 255, Paris, 75006 France; Université Paris-Descartes Paris 5, Faculté de Médecine, Paris, 75006 France; and Université Pierre et Marie Curie Paris 6, Institut des Cordeliers, Paris, 75006 France. <sup>2</sup>Institute for Genomics and Bioinformatics, Graz University of Technology, Graz, Austria. <sup>3</sup>Department of Pathology, <sup>4</sup>Department of Digestive Surgery, Avicenne Hospital, Bobigny, 93017 France. <sup>5</sup>Department of General and Digestive Surgery, <sup>6</sup>Department of Pathology, <sup>7</sup>Department of Immunology, Georges Pompidou European Hospital, AP-HP, Paris, 75015 France.

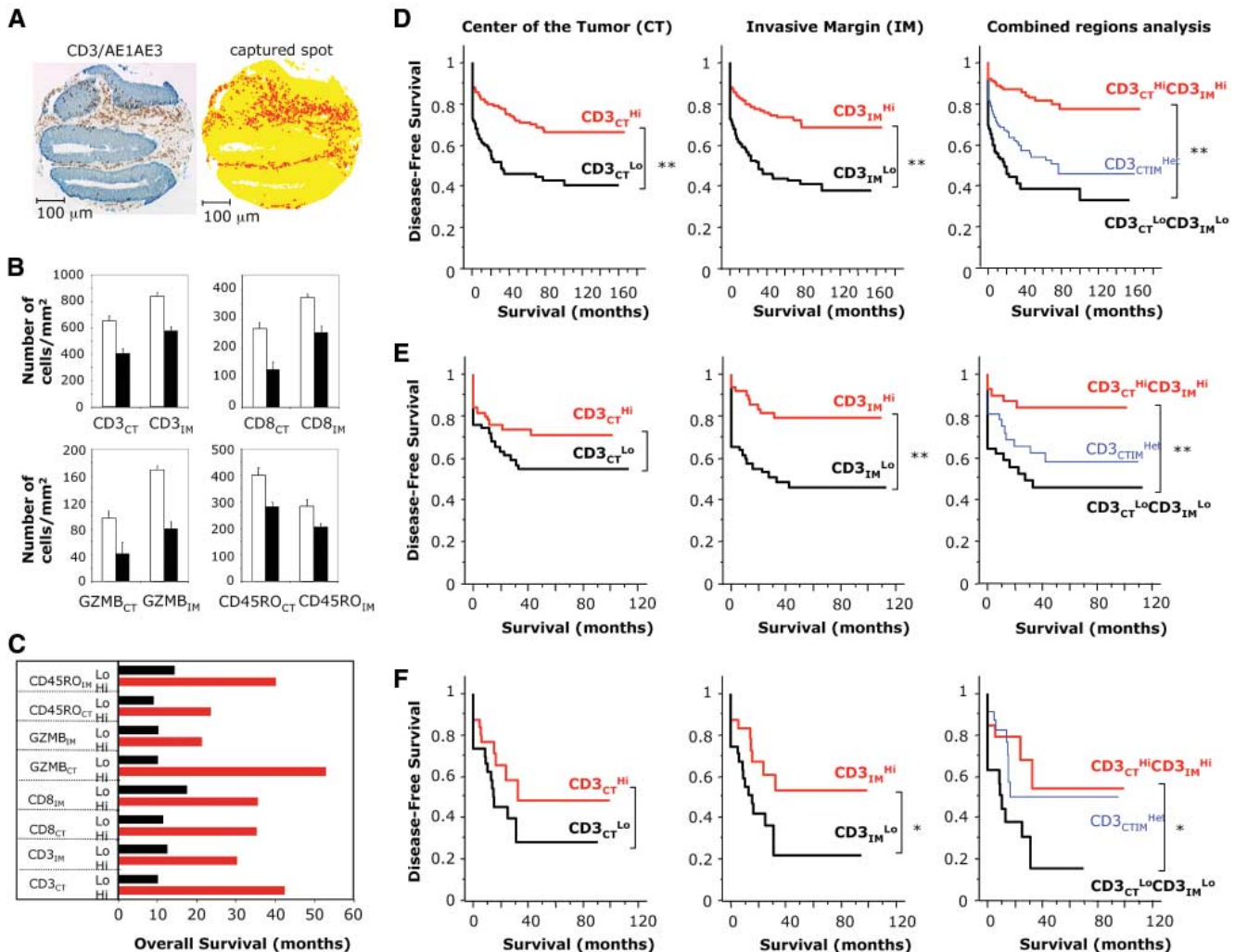
\*To whom correspondence should be addressed. E-mail: jerome.galon@u255.bhdc.jussieu.fr  
 †These authors contributed equally to this work.

(24) (100 repetitions for each marker), showing highly significant differences (tables S3 and S4).  $CD3_{CT}/CD3_{IM}$  density was associated with the smallest *P* values for disease-free and overall survival analyses ( $P = 7.6 \times 10^{-8}$  and  $P = 4.0 \times 10^{-7}$ , respectively) (tables S3 and S4). To confirm these results, we analyzed an additional cohort of patients who were different from those in the first series and a third cohort of CRC patients from another hospital. For each cohort, we determined the median cutoff values for  $CD3_{CT}/CD3_{IM}$  density (50% of patients with a high density and 50% of patients with a low density). The two independent cohorts (Fig. 2, E and F) confirmed the data

obtained on the first series (Fig. 2D). All statistical analyses were also performed for the subgroup of patients without concomitant distant metastasis [Union Internationale Centre le Cancer-Tumor Node Metastasis (UICC-TNM) cancer stages I, II, and III]. Significant *P* values were observed for  $CD3_{CT}/CD3_{IM}$ ,  $CD8_{CT}/CD8_{IM}$ , and  $CD45RO_{CT}/CD45RO_{IM}$  densities for predicting disease-free survival and overall survival (figs. S8 and S9 and tables S7 to S10).

We determined whether these immune criteria could discriminate patient outcome at each step of cancer progression. Patients were stratified according to the UICC-TNM classification (25) (Fig. 3A). A strong in situ immune reac-

tion in both tumor regions correlated with a favorable prognosis regardless of the local extent of the tumor and of invasion of regional lymph nodes (stages I, II, and III). Conversely, a weak in situ immune reaction in both tumor regions correlated with a poor prognosis even in patients with minimal tumor invasion (stage I) (Fig. 3B). We recently demonstrated the importance of the density of  $CD45RO^+$  memory T cells in limiting the tumor dissemination of CRCs (22). We found that patients with low densities of  $CD3^+$  cells and  $CD45RO^+$  memory T cells in both tumor regions (CT and IM) had a very poor prognosis, similar to that of patients with concomitant distant metastasis



**Fig. 2.** (A) (Left) A representative example of CD3 immunostaining of a CRC tissue microarray (top).  $CD3^+$  T cells (brown) and tumor cells (blue) are shown. (Right) Digital image analyzed with the image software SpotBrowser, with tissue represented in yellow and  $CD3^+$  cells represented in red. The densities of adaptive immune cells ( $CD3^+$ ,  $CD8^+$ ,  $GZMB^+$ , and  $CD45RO^+$  cells) were recorded as the number of positive cells per unit of tissue surface area. (B) Comparison of the mean ( $\pm$ SE) of immune cell densities in the CT and IM from patients with tumor recurrence (black bars) or without tumor recurrence (white bars). (C) Overall survival time for all patients, accounting for censoring (75th percentile), with high densities (red bars) or low densities (black bars) of adaptive immune cells in each tumor region (CT or IM). (D to F) Three independent cohorts of CRC patients were analyzed in

a blinded manner for  $CD3_{CT}/CD3_{IM}$  patterns [(D),  $n = 415$ ; (E),  $n = 119$ ; (F),  $n = 69$  patients]. Kaplan-Meier curves illustrate the duration of disease-free survival according to the  $CD3^+$  cell density in a single tumor region in the CT (left panels) or IM (middle panels) and in both tumor regions (right panels). In each cohort, for each tumor region, high (Hi) and low (Lo)  $CD3$  densities were plotted according to the cutoff value of  $CD3^+$  cell density defined at the median of the cohort (50% of patients with high cell density and 50% of patients with low cell density). In single-region analysis (left and middle panels), red lines indicate  $CD3^{Hi}$  and black lines indicate  $CD3^{Lo}$ . In combined analysis (right panels), red lines indicate  $CD3_{CT}^{Hi}CD3_{IM}^{Hi}$ , black lines indicate  $CD3_{CT}^{Lo}CD3_{IM}^{Lo}$ , and blue lines indicate heterogeneous  $CD3$  densities with  $CD3_{CT}^{Lo}$  plus  $CD3_{IM}^{Hi}$  or  $CD3_{CT}^{Hi}$  plus  $CD3_{IM}^{Lo}$  ( $CD3_{CTIM}^{Het}$ ).

(stage IV) (Fig. 3C). In multivariate analysis, after adjusting for tumor invasion (T stage), tumor differentiation, and lymph node invasion (N stage), CD3<sub>CT</sub>/CD3<sub>IM</sub> density (HiHi, Heterogeneous, and LoLo) remained an independent

prognostic factor, with the highest hazard ratio (HR) and the smallest *P* value in disease-free survival analysis [HR = 2.379; *P* = 1.4 × 10<sup>-6</sup>, corrected after (26)] (table S11). CD3<sub>CT</sub>/CD3<sub>IM</sub> density was the only independent parameter

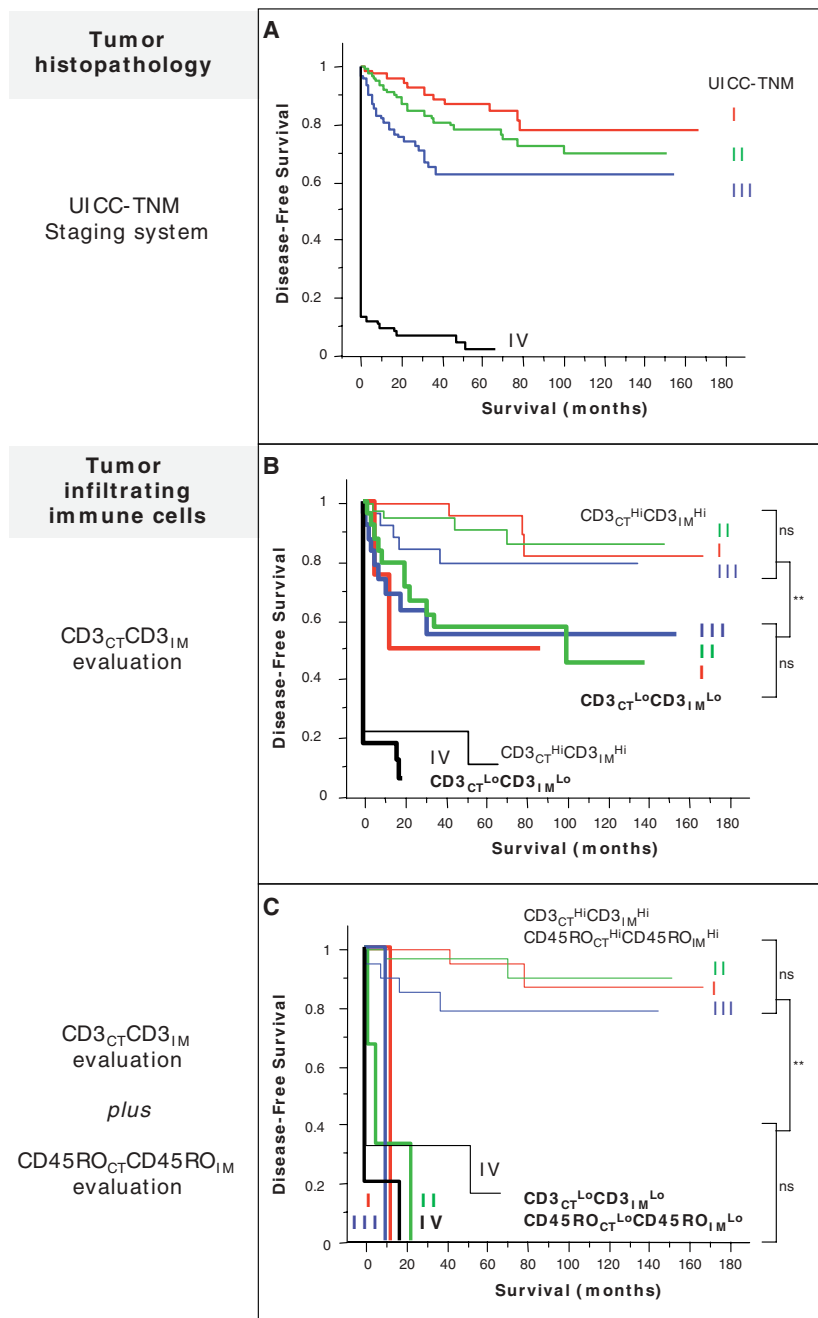
associated with overall survival (HR = 1.89; *P* = 1.2 × 10<sup>-5</sup>) (table S12). The histopathological parameters were no longer associated with disease-free and overall survival in patients with coordinated high or low densities of the immune markers in both tumor regions (HiHi versus LoLo) (tables S11 and S12).

Our results suggest that once human CRCs become clinically detectable, the adaptive immune response plays a role in preventing tumor recurrence. Despite immunoeediting (2), the beneficial effect of the adaptive immunity may persist throughout tumor progression (stages II and III). Intratumoral T cells could modify tumor stroma or tumor cells in ways that attenuate the metastatic potential of tumor cells. We found a positive correlation between the presence of markers for T<sub>H</sub>1 polarization and of cytotoxic and memory T cells and a low incidence of tumor recurrence. This argues for immune-mediated rejection of persistent tumor cells after surgery. We hypothesize that the trafficking properties and long-lasting anti-tumor capacity of memory T cells (27) play a central role in the control of tumor recurrence.

The type, density, and location of immune cells in CRCs had a prognostic value that was superior to and independent of those of the UICC-TNM classification (25). This suggests that time to recurrence and overall survival time are governed in large part by the state of the local adaptive immune response. The immunological criteria that we have used may lead to revision of the current indicators of clinical outcome and may help identify the high-risk patients who would benefit most from adjuvant therapy. Finally, this approach may be useful for the investigation of other tumor types.

**References and Notes**

1. G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, R. D. Schreiber, *Nat. Immunol.* **3**, 991 (2002).
2. G. P. Dunn, L. J. Old, R. D. Schreiber, *Annu. Rev. Immunol.* **22**, 329 (2004).
3. V. Shankaran *et al.*, *Nature* **410**, 1107 (2001).
4. G. Zhou, Z. Lu, J. D. McCadden, H. I. Levitsky, A. L. Marson, *J. Exp. Med.* **200**, 1581 (2004).
5. T. Boon, J. C. Cerottini, B. Van den Eynde, P. van der Bruggen, A. Van Pel, *Annu. Rev. Immunol.* **12**, 337 (1994).
6. P. van der Bruggen *et al.*, *Science* **254**, 1643 (1991).
7. S. J. Szabo, B. M. Sullivan, S. L. Peng, L. H. Glimcher, *Annu. Rev. Immunol.* **21**, 713 (2003).
8. K. E. de Visser, A. Eichten, L. M. Coussens, *Nat. Rev. Cancer* **6**, 24 (2006).
9. L. M. Coussens, Z. Werb, *Nature* **420**, 860 (2002).
10. F. R. Greten *et al.*, *Cell* **118**, 285 (2004).
11. D. Pardoll, *Nature* **411**, 1010 (2001).
12. D. Pardoll, *Annu. Rev. Immunol.* **21**, 807 (2003).
13. E. Pikarsky *et al.*, *Nature* **431**, 461 (2004).
14. J. W. Pollard, *Nat. Rev. Cancer* **4**, 71 (2004).
15. C. G. Clemente *et al.*, *Cancer* **77**, 1303 (1996).
16. P. K. Baier *et al.*, *Tumour Biol.* **19**, 205 (1998).
17. A. C. Diederichsen, J. B. Hjelmberg, P. B. Christensen, J. Zeuthen, C. Fenger, *Cancer Immunol. Immunother.* **52**, 423 (2003).
18. Y. Naito *et al.*, *Cancer Res.* **58**, 3491 (1998).
19. E. Sato *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18538 (2005).
20. L. Zhang *et al.*, *N. Engl. J. Med.* **348**, 203 (2003).



**Fig. 3.** (A) Kaplan-Meier curves illustrate the duration of disease-free survival according to the UICC-TNM stages [stage I, red line (*n* = 75 patients); stage II, green line (*n* = 137); stage III, blue line (*n* = 99), and stage IV, black line (*n* = 95)] in patients with CRCs. (B) Kaplan-Meier curves illustrate the duration of disease-free survival according to the UICC-TNM stages [as in (A)] and according to the density of CD3<sup>+</sup> cells in combined tumor regions (CD3<sub>CT</sub><sup>Lo</sup>CD3<sub>IM</sub><sup>Lo</sup>, thick lines, *n* = 93 patients; CD3<sub>CT</sub><sup>Hi</sup>CD3<sub>IM</sub><sup>Hi</sup>, thin lines, *n* = 109). The subgroup of patients that did not appear to have a coordinated *in situ* immune reaction in tumor regions (Hi/Lo or Lo/Hi for CD3<sup>+</sup> cell densities) presented similar Kaplan-Meier curves as the entire cohort (fig. S10). (C) Kaplan-Meier curves illustrate the duration of disease-free survival according to the UICC-TNM stages and to the density of CD3<sup>+</sup> and CD45RO<sup>+</sup> cells in combined tumor regions (CD3<sub>CT</sub><sup>Lo</sup>CD3<sub>IM</sub><sup>Lo</sup> plus CD45RO<sub>CT</sub><sup>Lo</sup>CD45RO<sub>IM</sub><sup>Lo</sup>, thick lines, *n* = 16 patients; CD3<sub>CT</sub><sup>Hi</sup>CD3<sub>IM</sub><sup>Hi</sup> plus CD45RO<sub>CT</sub><sup>Hi</sup>CD45RO<sub>IM</sub><sup>Hi</sup>, thin lines, *n* = 88). Cutoff values were 250, 640, 60, and 190 for CD3<sub>CT</sub>, CD3<sub>IM</sub>, CD45RO<sub>CT</sub>, and CD45RO<sub>IM</sub>, respectively. In (B) and (C), log-rank statistical test, \*\* *P* < 10<sup>-4</sup>; ns, not significant.

21. F. Sallusto, J. Geginat, A. Lanzavecchia, *Annu. Rev. Immunol.* **22**, 745 (2004).
22. F. Pages *et al.*, *N. Engl. J. Med.* **353**, 2654 (2005).
23. D. G. Altman, B. Lausen, W. Sauerbrei, M. Schumacher, *J. Natl. Cancer Inst.* **86**, 829 (1994).
24. D. Faraggi, R. Simon, *Stat. Med.* **15**, 2203 (1996).
25. L. Sobin, C. Wittekind, *TNM Classification of Malignant Tumors* (Wiley-Liss, ed. 6, New York, 2002).
26. N. Hollander, W. Sauerbrei, M. Schumacher, *Stat. Med.* **23**, 1701 (2004).
27. R. Xiang, H. N. Lode, S. D. Gillies, R. A. Reisfeld, *J. Immunol.* **163**, 3676 (1999).
28. J. Galon *et al.*, *FASEB J.* **16**, 61 (2002).
29. F. Pages *et al.*, *Blood* **105**, 1632 (2005).
30. A. Sturn, J. Quackenbush, Z. Trajanoski, *Bioinformatics* **18**, 207 (2002).
31. We thank A. Rodi and M. Pelegrin for expert technical assistance; I. Gresser for helpful comments and critical review of the manuscript; and D. Frucht, C. Anderson, and T. Pokrovskaya for critically reading the manuscript. Supported by the Association pour la Recherche sur le Cancer (ARC) through the Alliance pour la Recherche sur le Cancer network (ARECA), INSERM, Action Concertée Incitative ACI IMPBio, Université Paris-  
Descartes Paris 5, and the Austrian Federal Ministry of Education, Science and Culture (GBN-AU project Bioinformatics Integration Network).

### Supporting Online Material

www.sciencemag.org/cgi/content/full/313/5795/1960/DC1  
Materials and Methods  
Figs. S1 to S10  
Tables S1 to S12

25 April 2006; accepted 30 August 2006  
10.1126/science.1129139

## Volatile Chemical Cues Guide Host Location and Host Selection by Parasitic Plants

Justin B. Runyon, Mark C. Mescher, Consuelo M. De Moraes\*

The importance of plant volatiles in mediating interactions between plant species is much debated. Here, we demonstrate that the parasitic plant *Cuscuta pentagona* (dodder) uses volatile cues for host location. *Cuscuta pentagona* seedlings exhibit directed growth toward nearby tomato plants (*Lycopersicon esculentum*) and toward extracted tomato-plant volatiles presented in the absence of other cues. Impatiens (*Impatiens wallerana*) and wheat plants (*Triticum aestivum*) also elicit directed growth. Moreover, seedlings can distinguish tomato and wheat volatiles and preferentially grow toward the former. Several individual compounds from tomato and wheat elicit directed growth by *C. pentagona*, whereas one compound from wheat is repellent. These findings provide compelling evidence that volatiles mediate important ecological interactions among plant species.

Plant volatiles serve as important foraging cues for both insect herbivores and their natural enemies and can convey complex information regarding plant location, identity, and condition (1–5). It has been suggested that volatiles may have similar importance for interactions among plants, but such claims have remained controversial (6–13) and where plant-plant volatile effects have been demonstrated, their ecological importance remains unclear (6–9). Previous work on volatile-mediated interactions among plant species has dealt with the role of volatiles induced by herbivory or other environmental stressors in initiating defensive responses in neighboring plants (7, 14–19). Parasitic plants, which to survive must rapidly locate and attach to other plants, provide an alternative system in which host-plant volatiles might be expected to play an important role.

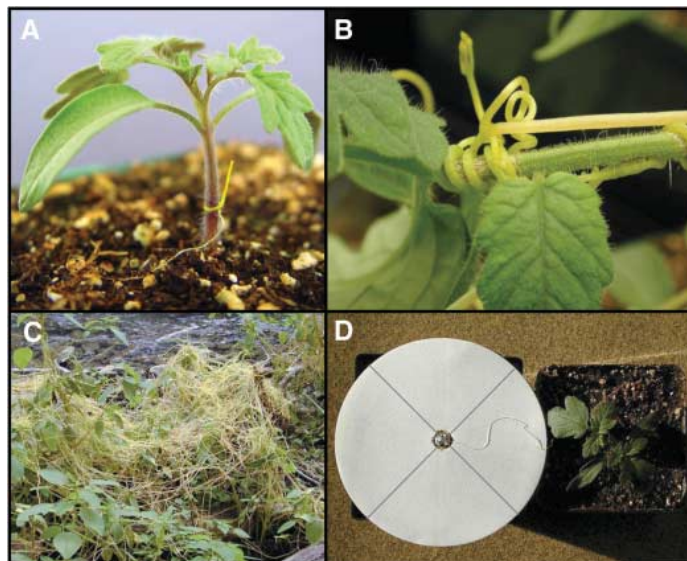
Parasitic plants are important components of both natural and agricultural ecosystems and have considerable influence on the structure and dynamics of the communities they inhabit (20, 21). Yet, little is known about the ecology of interactions between parasitic plants and their hosts. Like insect herbivores, para-

sitic plants exhibit various “foraging” patterns (22–25) and are capable of “selecting” among potential hosts (22–25), but the mechanisms involved in host location and discrimination are not well understood.

Flowering plants in the genus *Cuscuta* are obligate parasites with little photosynthetic capability; they obtain nutrients by attaching to aboveground shoots of other plants (26) (Fig. 1). *Cuscuta* spp. are important agricultural pests, included on the U.S. Department

of Agriculture’s *Top Ten Weeds List*, and can be difficult to control without also impacting host plants (27). Seeds of *Cuscuta* spp. contain minimal energy reserves, allowing growth of only several centimeters, and upon germination, the rootless seedlings must locate and attach to a suitable host within a few days (26). In some parasitic plants, contact with chemical cues secreted from host-plant roots is required for germination (28, 29), but *Cuscuta* spp. have no specialized germination requirements and must depend on seedling “foraging” for host-plant location (26) (fig. S1). After germination, *C. pentagona* seedlings exhibit a rotational growth habit (circumnutation) until contacting a host (26) (movie S1). Host secondary metabolites are known to influence the belowground growth of parasitic plants that attach to host roots (28, 29), and host-derived chemicals also induce haustorial development by these parasites (30). However, the role of host-derived compounds in aboveground host location by *Cuscuta* spp. has not previously been determined.

In this study, we explored host finding by seedlings of *C. pentagona*. First, we examined whether *C. pentagona* seedlings exhibit directed growth toward host plants (potted 20-day-old tomato seedlings). The basal end of a *C. pentagona* seedling was inserted into a water vial placed at the center of a dry filter-



**Fig. 1.** Parasitic plants in the genus *Cuscuta*. (A) *C. pentagona* seedling attaching to a tomato plant. (B) Vines of *C. pentagona* coiled around the petiole of a tomato leaf. (C) Growth habit of *Cuscuta*. (D) *C. pentagona* seedling growing toward a tomato plant across a filter-paper disc.

Department of Entomology, Pennsylvania State University, University Park, PA 16802, USA.

\*To whom correspondence should be addressed. E-mail: czd10@psu.edu

## Article 2

# "Coordination of Intratumoral Immune Reaction and Human Colorectal Cancer Recurrence"

Matthieu Camus, Marie Tosolini, Bernhard Mlecnik, Franck Pagès, Amos Kirilovsky, Anne Berger,  
Anne Costes, Gabriela Bindea, Pornpimol Charoentong, Patrick Bruneval, Zlatko Trajanoski,  
Wolf-Herman Fridman, and Jérôme Galon

Cancer Research, 2009, volume 69, numero 6, pages 2685-93

Nous avons précédemment démontré l'importance des paramètres immunologiques pour la prédiction très précise du devenir clinique des patients atteints de cancers colorectaux. Nous avons également mis en évidence que l'une des activités antitumorales du système immunitaire consiste à limiter la dissémination métastatique précoce des embolus tumoraux (VELIPI) au-delà du site tumoral primaire [19]. De plus, de fortes densités de lymphocytes T mémoires participant aux réponses immunitaires de types Th1, localisées au coeur des tumeurs et au niveau du front de progression tumorale, identifient les patients de pronostic favorable, quel que soit leur stade clinique (stades I, II et III). Cependant, la détection de métastase(s) reste un élément central dans la prise de décision d'administration des traitements adjuvants aux patients atteints de cancers colorectaux. Nous avons donc analysé le type de réponse immunitaire et les critères tumoraux (angiogenèse, invasion tumorale locale ou métastatique, immunosuppression) associés à une forte infiltration intratumorale en lymphocyte T chez les patients présentant ou non des métastases (patients META- ou META+). Nous avons ensuite déterminé la valeur pronostique de ces critères immunologiques et tumoraux en rapport avec le statut métastatique des patients. Dans cette étude, nous avons exploré la nature des cellules immunitaires présentes au site tumoral primaire ainsi que le degré d'organisation fonctionnelle de ces cellules traduisant l'efficacité des réponses immunitaires antitumorales en fonction de l'amplitude des infiltrats immunitaires et du statut métastatique des patients atteints de cancer colorectaux.

J'ai effectué les marquages des cellules immunitaires régulatrices (FoxP3), des macrophages (CD68) et des cellules dendritiques immatures (CD1a) par immunohistochimie sur les Tissue MicroArrays de 415 patients. J'ai également effectué les marquages des cellules en prolifération (Ki-67) et en apoptose (M30) en pleine lame sur 39 patients. J'ai analysé ces différents marquages à l'aide du logiciel Spotbrowser au centre et front de progression de la tumeur (figure supplémentaire 4 et données non montrées).

## Résumé

Le rôle du système immunitaire dans le contrôle de la progression tumorale a été établi dans des modèles murins. Toutefois, les effets des réponses immunitaires et de l'échappement tumoral sur le pronostic des patients dans le cadre d'un cancer humain sont peu compris.

Nous avons étudié les paramètres cellulaires et moléculaires qui pourraient décrire les réponses immunitaires *in situ* dans des cancers colorectaux humains en fonction de paramètres cliniques tel que l'envahissement de ganglions ou d'organes à distance (META- ou META+). Les cellules infiltrant les tumeurs primaires ont été analysées par cytométrie de flux et par RT-PCR quantitative.

L'analyse des corrélations du phénotypage effectué par cytométrie de flux de tumeurs sans métastase (META-) et avec un fort infiltrat lymphocytaire, montre différents clusters. Un cluster est associé à des cellules immunitaires innées (macrophages, cellules dendritiques, natural killer) et des cellules T en phase précoce d'activation. Un autre cluster est associé aux marqueurs de cytotoxicité et à la sous-population de cellules T effectrices mémoires. Ce profil est altéré dans les tumeurs avec un faible infiltrat et absent dans des tumeurs avec métastases (META+).

La co-expression du gène granulysine (cytotoxicité) et IRF1 (Th1) permet de prédire la survie du patient indépendamment du statut métastatique. Cependant l'expression du VEGF (vascularisation) réduit significativement la survie des patients exprimant de forts niveaux de granulysine.

Une réponse immunitaire intratumorale avec une coordination efficace est associée avec un pronostic favorable. Nous avons pu ainsi déterminer quatre groupes principaux de patients. Le premier groupe présente une forte densité de cellules T avec une bonne coordination immunitaire et un niveau faible

de VEGF. Leur risque de récurrence à 2 et 10 ans est identique et de 10%. Le deuxième et le troisième groupe ont une forte densité de cellules T mais soit une perte de la coordination de la réponse immunitaire, soit un niveau élevé de VEGF. Leur risque de récurrence à 2 ans est alors de 50%. Le quatrième groupe a le risque de récurrence sur 2 ans le plus élevé (80%). Leurs densités de lymphocytes T sont basses et ils n'ont pas de coordination de la réponse immunitaire.



# Coordination of Intratumoral Immune Reaction and Human Colorectal Cancer Recurrence

Matthieu Camus,<sup>1,2,3</sup> Marie Tosolini,<sup>1,2,3</sup> Bernhard Mlecnik,<sup>1,2,3</sup> Franck Pagès,<sup>1,2,3,4</sup> Amos Kirilovsky,<sup>1,2,3</sup> Anne Berger,<sup>5</sup> Anne Costes,<sup>1,2,3</sup> Gabriela Bindea,<sup>1,2,3,7</sup> Pornpimol Charoentong,<sup>7</sup> Patrick Bruneval,<sup>6</sup> Zlatko Trajanoski,<sup>7</sup> Wolf-Herman Fridman,<sup>1,2,3,4</sup> and Jérôme Galon<sup>1,2,3</sup>

<sup>1</sup>Integrative Cancer Immunology INSERM AVENIR Team 15, INSERM U872; <sup>2</sup>Cordeliers Research Centre, Université Pierre et Marie Curie Paris 6; <sup>3</sup>Université Paris-Descartes; Departments of <sup>4</sup>Immunology, <sup>5</sup>General and Digestive Surgery, and <sup>6</sup>Pathology, Georges Pompidou European Hospital, Paris, France; and <sup>7</sup>Institute for Genomics and Bioinformatics, Graz University of Technology, Graz, Austria

## Abstract

**A role for the immune system in controlling the progression of solid tumors has been established in several mouse models. However, the effect of immune responses and tumor escape on patient prognosis in the context of human cancer is poorly understood. Here, we investigate the cellular and molecular parameters that could describe *in situ* immune responses in human colorectal cancer according to clinical parameters of metastatic lymph node or distant organ invasion (META– or META+ patients). Primary tumor samples of colorectal carcinoma were analyzed by integrating large-scale phenotypic (flow cytometry, 39 patients) and gene expression (real time reverse transcription-PCR, 103 patients) data sets related to immune and protumoral processes. In META– colorectal cancer primary tumors with high densities of T cells, we observed significant positive correlations between markers of innate immune cells [tumor-associated macrophages, dendritic cells, natural killer (NK) cells, and NKT cells] and markers of early-activated T cells. Significant correlations were also observed between markers of cytotoxic and effector memory T-cell subpopulations. These correlation profiles were absent in tumors with low T-cell infiltrates and were altered in META+ tumors with high T-cell infiltrates. We show that the coexpression of genes mediating cytotoxicity (*GNLY*) and Th1 adaptive immune responses (*IRF1*) accurately predicted patient survival independently of the metastatic status. High intratumoral mRNA expression of the proangiogenic mediator vascular endothelial growth factor was associated with significantly reduced survival rates in patients expressing high mRNA levels of *GNLY*. Investigation of the colorectal cancer primary tumor microenvironment allowed us to uncover the association of favorable outcomes with efficient coordination of the intratumoral immune response.** [Cancer Res 2009;69(6):2685–93]

## Introduction

Cancer progression is a complex process involving host-tumor interactions through multiple molecular and cellular factors of the tumor microenvironment (1). Tumors may be vulnerable to

immune destruction. As revealed by experiments in immune-deficient mice, immune responses mediated by IFN $\gamma$  (2, 3) and cytotoxic mediators such as perforin (4, 5) secreted by lymphocytes are involved in cancer immunosurveillance (6, 7). In human cancer, complex tumor-host interactions are less well documented. However, lymphocytes were also shown to participate in anti-tumoral responses (8). Consistent with findings in melanoma (9) and ovarian cancer (10, 11), tumor-infiltrating T cells were associated with improved clinical outcome and survival in colorectal cancer patients (12–16).

We recently highlighted intratumoral memory T cells as the major immune effector cells significantly associated with the decrease of early metastatic events (tumor emboli) and the prevention of relapse in colorectal cancer patients (17). Furthermore, we revealed the importance to patient prognosis of the nature, functional orientation, density, and localization of immune cell populations within the primary tumor. Multivariate Cox analysis showed that immune patterns remained the unique parameter significantly associated with prognosis, whereas T stage, N stage, and differentiation of the tumor were not significant when adjusted to immune patterns (18). Patients with cancers at nonmetastatic stages had prognoses as bleak as patients with metastatic tumors, if presenting a low intratumoral adaptive immune reaction. Conversely, patients with metastatic tumors eliciting a high intratumoral immune reaction were of better prognosis. Thus, the amplitude of adaptive immune reaction within the primary tumor was a better predictor of survival than traditional clinical parameters (19).

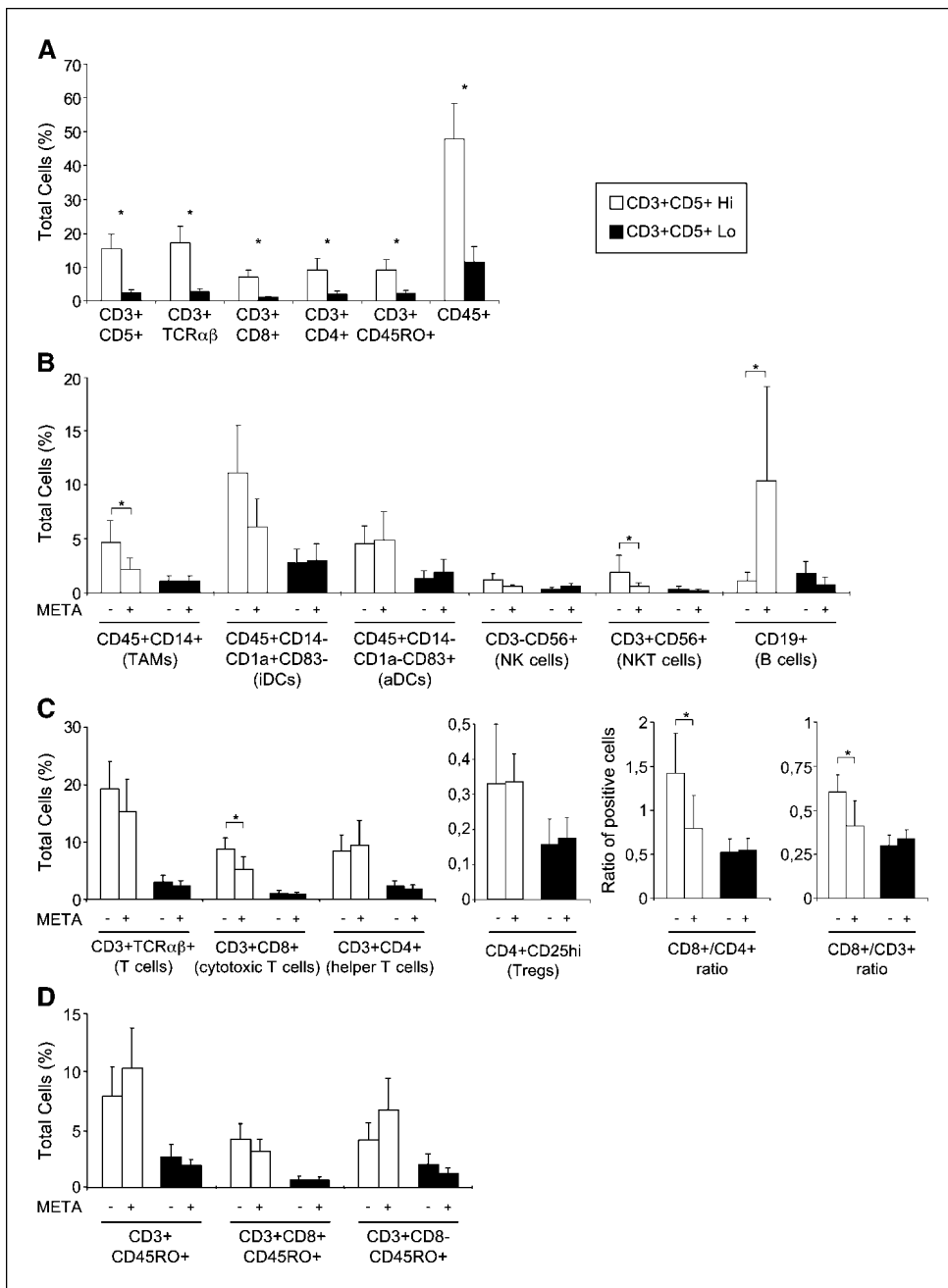
However, the intrinsic capability of tumor cells to promote their own development (20) may allow tumors to overwhelm immune system activity. For instance, angiogenesis mediated by vascular endothelial growth factor (VEGF) is critical to the growth (by providing oxygen and nutrients) and malignant dissemination (providing a route for metastases) of solid tumors (21, 22). Furthermore, under the pressure of antitumoral immune activity, selection and outgrowth of variant tumor cells with reduced immunogenicity could occur (8, 22, 23). Thus, during cancer progression, tumor cells may acquire immune tolerance mechanisms by generating complex immunosuppressive networks at the tumor site (24, 25) involving interleukin (IL)-10 and transforming growth factor  $\beta$  (TGF $\beta$ ; refs. 26, 27) as well as T-cell-specific coinhibitory molecules (CTLA-4 and PD-1; refs. 28, 29).

In this work, we attempted to describe in a comprehensive manner the immune reaction in primary colorectal tumors of patients with high or low densities of infiltrating T cells. Furthermore, we compared the immune microenvironment in patients presenting with invaded lymph nodes and/or distant metastases [META+ patients: Union Internationale Contre le

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

**Requests for reprints:** Jérôme Galon, AVENIR Team 15, Cordeliers Research Centre, INSERM U872, 15 rue de L'École de Médecine, 75006 Paris, France. Phone: 33-1-5310-0404; Fax: 33-1-4051-0420; E-mail: [jerome.galon@crc.jussieu.fr](mailto:jerome.galon@crc.jussieu.fr).

©2009 American Association for Cancer Research.  
doi:10.1158/0008-5472.CAN-08-2654



**Figure 1.** Immune cell populations within primary colorectal tumors. Patients ( $n = 39$ ) were classified according to the mean percentage of CD3<sup>+</sup>CD5<sup>+</sup> cells among total cells within tumors (white columns, CD3<sup>+</sup>CD5<sup>+</sup>Hi; black columns, CD3<sup>+</sup>CD5<sup>+</sup>Lo) and the metastatic status (META- Hi,  $n = 6$ ; META- Lo,  $n = 10$ ; META+ Hi,  $n = 7$ ; META+ Lo,  $n = 16$ ). Cell populations were represented as the mean percentage of positive cells; bars, SE. \*,  $P < 0.05$ , Mann-Whitney test. A, T cells (CD3<sup>+</sup>CD5<sup>+</sup>, CD3<sup>+</sup>TCRαβ<sup>+</sup>), cytotoxic T cells (CD3<sup>+</sup>CD8<sup>+</sup>), helper T cells (CD3<sup>+</sup>CD4<sup>+</sup>), memory T cells (CD3<sup>+</sup>CD45RO<sup>+</sup>), and lymphoid cells (CD45<sup>+</sup>). B, tumor-associated macrophages (TAMs; CD45<sup>+</sup>CD14<sup>+</sup>), immature dendritic cells (iDCs; CD45<sup>+</sup>CD14<sup>+</sup>CD1a<sup>+</sup>CD83<sup>-</sup>), activated dendritic cells (aDCs; CD45<sup>+</sup>CD14<sup>-</sup>CD1a<sup>-</sup>CD83<sup>+</sup>), NK cells (CD3<sup>-</sup>CD56<sup>+</sup>), NKT cells (CD3<sup>+</sup>CD56<sup>+</sup>), and B cells (CD19<sup>+</sup>). C, left, T cells (CD3<sup>+</sup>TCRαβ<sup>+</sup>), cytotoxic T cells (CD3<sup>+</sup>CD8<sup>+</sup>), helper T cells (CD3<sup>+</sup>CD4<sup>+</sup>), and regulatory T cells (Tregs; CD4<sup>+</sup>CD25<sup>hi</sup>). Right, ratios of CD8<sup>+</sup>/CD4<sup>+</sup> and CD8<sup>+</sup>/CD3<sup>+</sup> cell subpopulations. D, memory T cells (CD3<sup>+</sup>CD45RO<sup>+</sup>), cytotoxic memory T cells (CD3<sup>+</sup>CD8<sup>+</sup>CD45RO<sup>+</sup>), and helper memory T cells (CD3<sup>+</sup>CD4<sup>+</sup>CD45RO<sup>+</sup>).

Cancer (UICC) tumor-node-metastasis (TNM) stages III–IV) or without such metastases (META– patients: UICC-TNM stages I–II; ref. 30). We analyzed immune cell phenotypic clusters, or “phenoclusters” (31), obtained by grouping markers according to similar levels of expression. This allowed us to uncover functional marker patterns of efficient and coordinated antitumoral immune responses that represent powerful prognostic criteria for colorectal cancer clinical outcome. At the cellular level, a high degree of functional coordination between intratumoral immune cells could be observed at the primary tumor sites of both META– and META+ colorectal cancer patients. At the molecular level, the coexpression of genes related to the Th1 immune response [IFN-regulatory factor 1 (*IRF1*)] and cytotoxicity [granulysin (*GNLY*)] had strong prognostic values. Finally, we studied several tumor-promoting mechanisms including immunosuppression, angiogen-

esis, tumor survival, and local and metastatic invasion. Analysis of *in situ* gene expression of protumoral markers in combination with immune parameters revealed that angiogenesis (*VEGF*) was associated with increased risks of colorectal cancer relapse in patients nonetheless presenting evidence of strong intratumoral immune responses.

## Materials and Methods

All details about Materials and Methods are available online.

**Patients and database.** Patients with colorectal cancer ( $n = 566$ ) who underwent a primary resection at the Laennec/HEGP Hospital between 1986 and 2004 were randomly selected. Time to recurrence or disease-free time was defined as the time period from the date of surgery to confirmed tumor relapse date for relapsed patients and from the date of surgery to the date of last follow-up for disease-free patients.

**Large-scale flow cytometric analysis.** Cells were extracted from 39 fresh tumors, resuspended in PBS/0.5% bovine serum albumin and incubated for 30 min at 4°C with antibodies against immune cell markers for large-scale phenotypic analysis of T cells and with relevant isotype controls. Analyses were done with a four-color fluorescence-activated cell sorter (FACSCalibur, Becton Dickinson) and CellQuest software (Becton Dickinson). Analyzed markers are presented in Supplementary Fig. S3. Complete-linkage hierarchical clustering was applied and the results were displayed with the use of the Genesis program (32, 33). Correlation matrices were constructed by calculation of Pearson correlation coefficients for all marker combinations, followed by unsupervised hierarchical clustering.

**Real-time reverse transcription-PCR assay.** Tissue samples were snap-frozen. Total RNA was extracted by homogenization with RNeasy isolation kit (Qiagen). The integrity and the quantity of the RNA were evaluated on Bioanalyzer-2100 (Agilent Technologies). Samples ( $n = 103$ ) were assessed for gene expression analysis of the following 17 genes (see details about gene expression and name in Supplementary data): *CD3ζ*, *CD4*, *CD8α*, *TBX21*, *IRF1*, *IFNγ*, *GNLY*, *GZMB*, *GATA3*, *FOXP3*, *CEACAM1*, *CEA*, *EBAG9*, *BIRC5*, *IL-10*, *TGFβ*, and *VEGF*. Quantitative real-time TaqMan PCR was done using Low-Density-Arrays and the 7900 robotic real-time PCR system (Applied Biosystems). 18S primers and probes were used as internal controls.

**Construction of tissue microarrays.** Using a tissue microarray instrument (Beecher Instruments, Alphelys), we removed two representative areas of the tumor (center and invasive margin from paraffin-embedded tissue blocks). Tissue microarrays were cut into 5-μm sections for immunohistochemical staining.

**Immunohistochemistry.** After antigen retrieval and quenching of endogenous peroxidase activity, sections were incubated for 60 min at room temperature with monoclonal antibodies against CD3 (SP7), CD8 (4B11), CD1a (O10), Ki67 (SP6; Neomarkers), CD68 (PGM1; DAKO), FoxP3 (ab20034; Abcam), and M30 cytoDEATH (Alexis Biochemicals). The

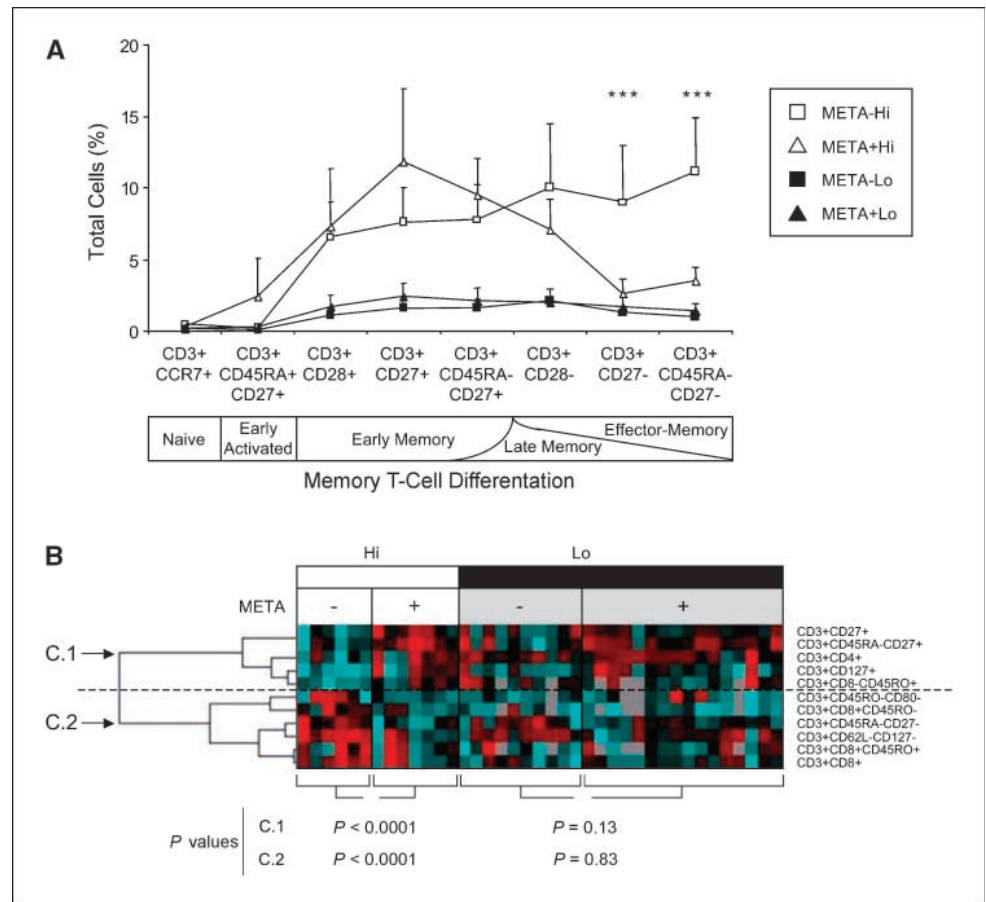
Envision+ system (enzyme-conjugated polymer backbone coupled to secondary antibodies) and 3,3'-diaminobenzidine chromogen were applied (DAKO). Tissue sections were counterstained with Harris's hematoxylin. Isotype-matched mouse monoclonal antibodies were used as negative controls.

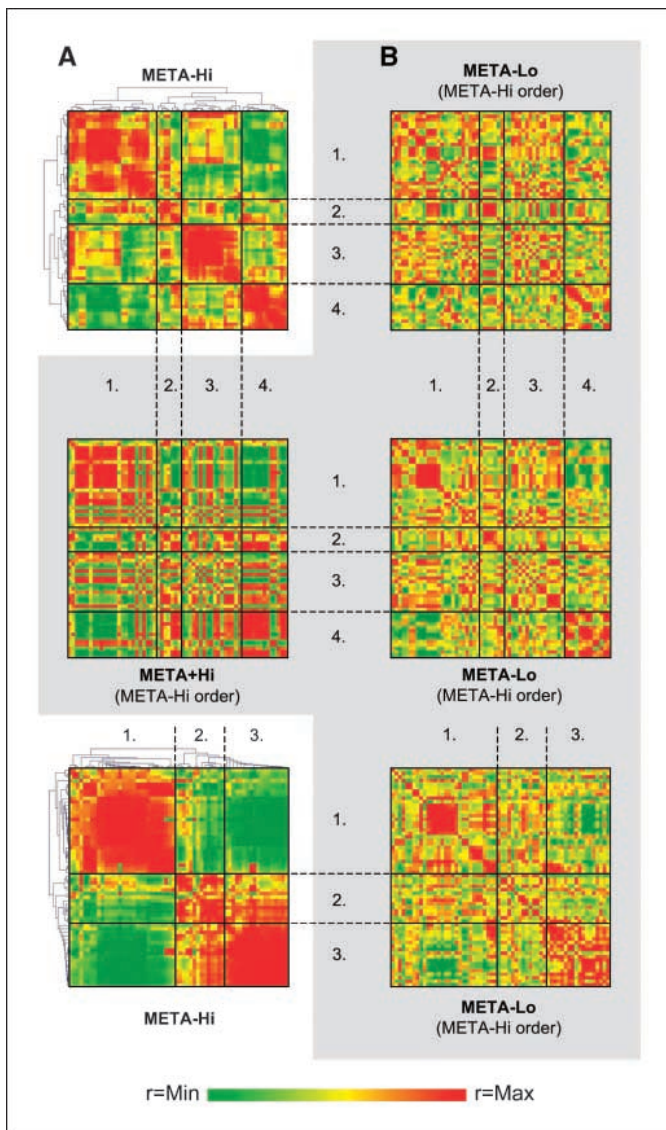
**Statistical analysis.** Kaplan-Meier curves were used to assess the influence of immune and tumoral parameters on disease-free survival. The significance of these parameters was assessed by univariate analysis with the use of the log-rank test. To identify markers with significant different levels of expression among tissues, Wilcoxon-Mann-Whitney and *t* tests (ANOVA) were used.  $P < 0.05$  was considered to indicate statistical significance. All analyses were done with the use of statistical software programs R and StatView.

**Results**

**Intratumoral distribution of immune cell populations.** We first investigated the immune cellular profiles of patients by flow cytometry with 39 freshly resected primary tumors. Percentages of positive cells for distinct markers were calculated among all cells (tumor and immune cells), thus reflecting the density of cells within tumors. Intratumoral T-cell densities were evaluated according to the mean percentage of double-positive cells for the T-cell-specific markers CD3 and CD5 among all samples. Patients presenting a percentage of CD3<sup>+</sup>CD5<sup>+</sup> T cells superior to this mean (6.7% of all cells) were named "Hi patients" and otherwise "Lo patients." As a control, the percentages of CD3<sup>+</sup>CD5<sup>+</sup> and CD3<sup>+</sup>TCRαβ<sup>+</sup> T cells were represented (Fig. 1A). In Hi patients, there were significant higher densities of T cells of both cytotoxic (CD3<sup>+</sup>CD8<sup>+</sup>) and helper (CD3<sup>+</sup>CD4<sup>+</sup>) phenotypes and of memory

**Figure 2.** T-cell populations within primary colorectal tumors. **A**, T-cell memory differentiation markers (white squares, META- Hi; white triangles, META+ Hi; black squares, META- Lo; black triangles, META+ Lo). Cell populations were represented as the mean percentage of positive cells; bars, SE. *P* values (Mann-Whitney test) were presented in Supplementary Fig. S1 (\*,  $P < 0.05$ ). **B**, hierarchical clustering of 11 marker combinations among CD3<sup>+</sup> cells with significant differential expression among the four groups of patients ( $P < 0.05$ ). Combinations of surface markers were plotted from the minimal (blue) to the maximal (red) level of expression. Gray, not determined.





**Figure 3.** Correlation matrices of flow cytometry data. *P* values and Pearson correlation coefficients (*r*) were calculated between 62 marker combinations that were specific for T cells (markers in combination with CD3) and for major immune cell populations ("total" prefix), presented in Supplementary Fig. S2. *r* values were plotted from *r* = min (green) to *r* = max (red) in matrix representation, followed by unsupervised hierarchical clustering. Clustered markers were presented in Supplementary Fig. S3. Correlation matrices were independently clustered or arrayed according to the clustering of other correlation matrices (gray area). A, top, META– Hi patients; center and bottom, META+ Hi patients. B, top, META– Lo patients; center and bottom, META+ Lo patients.

phenotype (CD3<sup>+</sup>CD45RO<sup>+</sup>) compared with Lo patients (Fig. 1A). The distribution of global lymphoid cell populations (CD45<sup>+</sup>) was consistent with those of T cells (Fig. 1A).

We compared the distributions of the major intratumoral immune cell populations according to the metastatic status of the patients: META– patients, no metastases (stages I–II); META+ patients, metastases in lymph node (stage III) and/or distant organ (stage IV). In Lo patients, no differences were found between META– and META+ patients in the distribution of tumor-associated macrophages, immature dendritic cells, activated dendritic cells, natural killer (NK) cells, NKT cells, or B cells. In contrast, in Hi patients, significantly lower percentages of tumor-associated macrophages and NKT cells were observed in

META+ Hi patients compared with META– Hi patients. Conversely, B-cell density was significantly higher in META+ Hi patients compared with META– Hi patients (Fig. 1B). META+ Hi patients had significantly decreased densities of cytotoxic T cells (CD3<sup>+</sup>CD8<sup>+</sup>) compared with META– Hi patients, whereas no significant differences were observed for helper T cells (CD3<sup>+</sup>CD4<sup>+</sup>) or regulatory T cells (CD4<sup>+</sup>CD25<sup>hi</sup>; Fig. 1C, left). Finally, CD8<sup>+</sup>/CD4<sup>+</sup> and CD8<sup>+</sup>/CD3<sup>+</sup> cell ratios were significantly higher in META– Hi patients compared with META+ Hi and META+ Lo patients (Fig. 1C, right).

**Memory T-cell differentiation.** No differences in the distribution of CD45RO<sup>+</sup> memory T-cell subpopulations were observed among the patient groups (Fig. 1D). However, we more precisely assessed the density of T cells along memory differentiation steps based on the differential expression of CCR7, CD45RA, CD27, and CD28 markers by CD3<sup>+</sup> T cells (Fig. 2A). Very few naive (CCR7<sup>+</sup>) T cells were detected within primary tumors. In META– Lo (black squares) and META+ Lo patients (black triangles), despite low densities of T cells, similar levels of memory T-cell subpopulations from early (CD28<sup>+</sup>) to late (CD45RA<sup>–</sup>CD27<sup>–</sup>) memory were observed. META– Hi patients (white squares) presented high densities of all memory T-cell subpopulations. In contrast, META+ Hi patients (white triangles) presented a significant decrease in the densities of T cells at late stages of memory differentiation (CD27<sup>–</sup>, CD45RA<sup>–</sup>) with percentages comparable to Lo patients and significantly inferior to META– Hi patients.

Eleven marker combinations expressed among CD3<sup>+</sup> T cells were found significantly differentially expressed between META– Hi and META+ Hi patients. After hierarchical clustering of these markers, two major clusters (C.1 and C.2) were found (Fig. 2B). In C.1, CD4<sup>+</sup> T-cell subpopulation markers (CD3<sup>+</sup>CD4<sup>+</sup>) and related memory markers (CD3<sup>+</sup>CD8<sup>–</sup>CD45RO<sup>+</sup>) grouped with early memory T-cell markers (CD3<sup>+</sup>CD127<sup>+</sup>, CD3<sup>+</sup>CD27<sup>+</sup>, CD3<sup>+</sup>CD45RA<sup>–</sup>CD27<sup>+</sup>). In C.2, CD8<sup>+</sup> T-cell subpopulation markers (CD3<sup>+</sup>CD8<sup>+</sup>) and related memory/effector T-cell markers (CD8<sup>+</sup>CD45RO<sup>+/–</sup>) grouped with effector memory T-cell markers (CD3<sup>+</sup>CD45RA<sup>–</sup>CD27<sup>–</sup>) and final effector T-cell markers (CD3<sup>+</sup>CD45RO<sup>–</sup>, CD3<sup>+</sup>CD62L<sup>–</sup>CD127<sup>–</sup>). Whereas no distinct pattern was observed in Lo patients, META– Hi patients presented a significant increase of CD8/effector memory T-cell subpopulations (red squares; C.2) compared with META+ Hi patients that had a majority of CD4/early memory T cells (C.1). These observations suggested that complete memory T-cell differentiation was associated with a higher proportion of cytotoxic T cells within highly infiltrated tumors and preferentially occurred in META– Hi patients compared with META+ Hi patients.

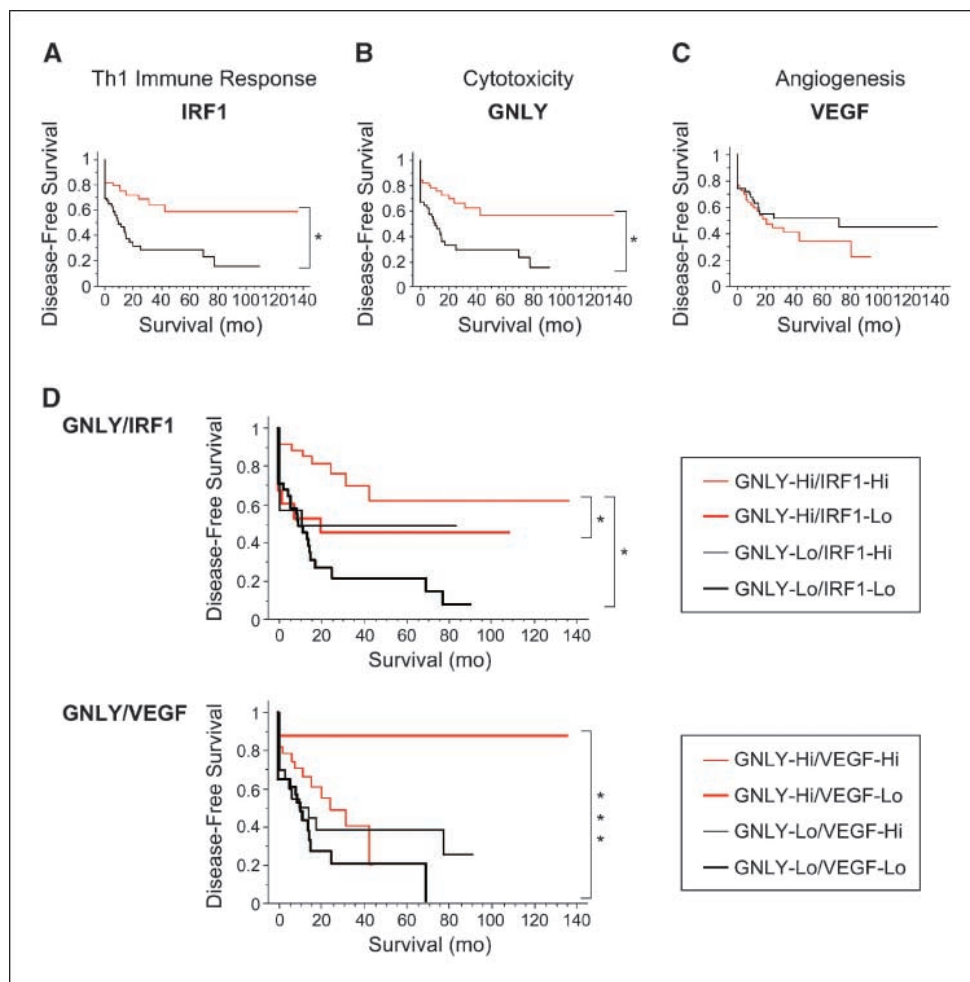
**Association between CD8 T cells and complete memory T-cell differentiation.** Evaluation of intratumoral immune coordination was assessed by analyzing the correlations between 62 combinations of cell surface markers of total intratumoral immune cell populations and T-cell subpopulations. For each patient group, pairwise comparisons of the markers were done by measuring Pearson correlation coefficients (*r*) and related *P* values (Supplementary Fig. S2). The relationships implied by these correlations were visualized by using unsupervised hierarchical clustering of *r* values (Fig. 3). The clustered markers were presented in Supplementary Fig. S3. Comparison of META– Hi patients (Fig. 3A, top) with other patients was assessed by the construction of META+ Hi (Fig. 3A, center), META– Lo (Fig. 3B, top), and META+ Lo (Fig. 3B, center) correlation matrices arrayed according to META– Hi matrix unsupervised clustering.

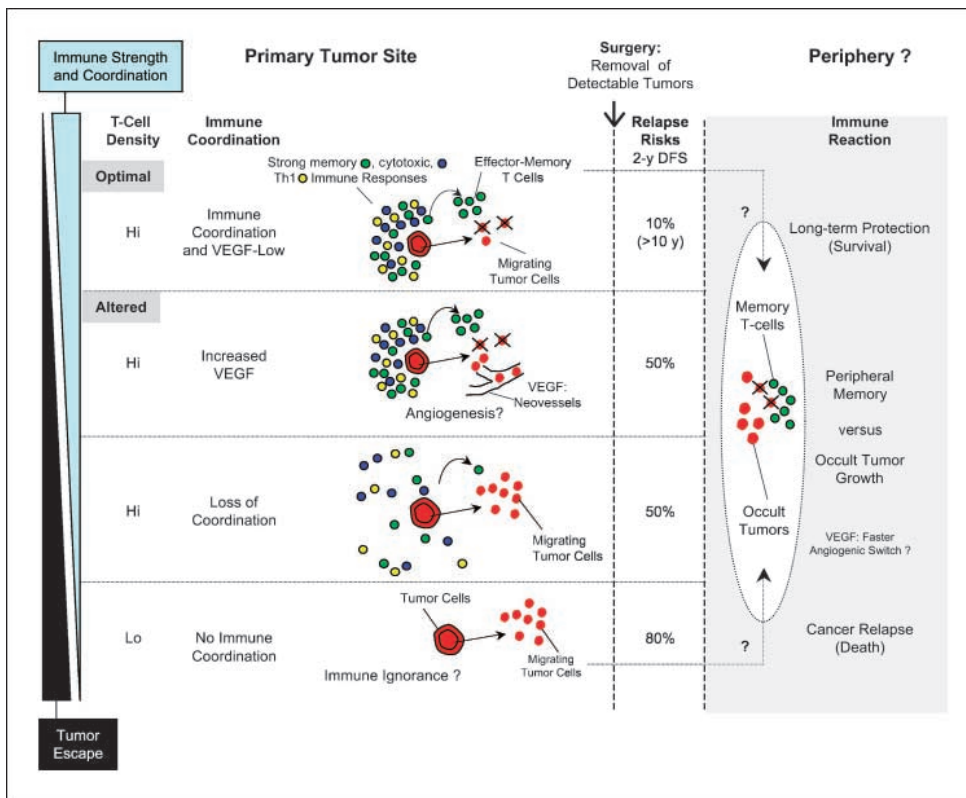
META- Hi patients displayed a correlation matrix with four major clusters (Fig. 3A, top). Cluster 1 contained markers of total T cells, CD4 T cells, B cells, NK cells, and activated dendritic cells, as well as CD4 T-cell subpopulation markers (CD4<sup>+</sup> or CD8<sup>-</sup> in combination with CD25<sup>+/-</sup>, CD26<sup>+/-</sup>, CD103<sup>+/-</sup>, CCR7<sup>-</sup>, CD45RO<sup>-</sup>) and early memory T-cell markers (CD28<sup>+</sup>, CD27<sup>+</sup>, CD45RA<sup>-</sup>CD27<sup>+</sup>). Significant positive correlations were found for the CD4 marker with CD27 and CD45RA<sup>-</sup>CD27<sup>+</sup> markers ( $P = 0.03$  for both correlations). Cluster 2 contained naive CCR7<sup>+</sup> T cells. Total CD8 T-cell marker from this cluster positively correlated with total T cells in cluster 1 ( $P = 0.03$ ). In cluster 4, CD8 T-cell subpopulation markers (CD8<sup>+</sup> or CD4<sup>-</sup> in combination with CD25<sup>+/-</sup>, CD26<sup>-</sup>, CD103<sup>+</sup>, CCR7<sup>-</sup>, CD45RO<sup>+</sup>) were grouped with markers of late-stage memory T-cell differentiation (CD45RA<sup>-</sup>CD27<sup>-</sup>CD127<sup>-</sup>CD62L<sup>-</sup>). Significant positive correlations were found between CD8 T-cell subpopulation markers and effector memory T-cell markers (CD45RA<sup>-</sup>CD27<sup>-</sup>/CD4<sup>-</sup>CD103<sup>+</sup>;  $P = 0.02$ ) and final effector T-cell markers (CD127<sup>-</sup>CD62L<sup>-</sup>/CD8<sup>+</sup>CCR7<sup>-</sup>;  $P = 0.02$ ). As a control, correlation between CD4 and CD8 T-cell subpopulation markers was negative (CD3<sup>+</sup>CD4<sup>+</sup>/CD3<sup>+</sup>CD8<sup>+</sup>;  $r = -0.871$ ,  $P = 0.02$ ). Interestingly, cluster 4 (CD8/effector memory T cells) and cluster 1 (CD4/early memory T cells) were, globally, strongly inversely correlated. In cluster 3, early differentiated (CD45RA<sup>+</sup>CD27<sup>+</sup>) and activated (CD25<sup>+</sup>, CD26<sup>+</sup>, CD69<sup>+</sup>) T-cell markers, as well as the global memory T-cell marker (CD45RO<sup>+</sup>), were grouped with markers of

innate immune cell populations: tumor-associated macrophages, immature dendritic cells, and NKT cells. Strong positive correlations were found between all markers of early T-cell activation (CD45RA<sup>+</sup>CD27<sup>+</sup>/CD25<sup>+</sup>/CD26<sup>+</sup>/CD69<sup>+</sup>;  $P < 0.05$  for all combinations) and the markers of tumor-associated macrophages and immature dendritic cells ( $P = 0.01$ ). These two functional groups of markers were positively correlated (tumor-associated macrophages/CD25<sup>+</sup>/CD26<sup>+</sup>/CD69<sup>+</sup> and immature dendritic cells/CD25<sup>+</sup>/CD26<sup>+</sup>;  $P < 0.05$  for all combinations). Thus, cluster 3 may illustrate the role of the innate immune compartment for T-cell priming, activation, and memory differentiation required for efficient adaptive immune responses.

In contrast, cluster 3 was entirely disrupted in META+ Hi patients (Fig. 3A, center). Indeed, the META+ Hi clustered correlation matrix (Fig. 3A, bottom) displayed two inversely correlated groups of clusters (cluster 1 versus clusters 2 and 3). As observed in the META- Hi matrix, cluster 1 in META+ Hi matrix contained the markers of CD4 and early memory T-cell subpopulations markers (CD4<sup>+</sup>/CD28<sup>+</sup>;  $P = 0.008$ ), as well as markers of total T cells, CD4 T cells, and B cells. In clusters 2 and 3, CD8 memory T-cell marker (CD8<sup>+</sup>CD45RO<sup>+</sup>) was significantly positively correlated with both effector memory (CD45RA<sup>-</sup>CD27<sup>-</sup>) and final effector (CD127<sup>-</sup>CD62L<sup>-</sup>) T-cell markers ( $P = 0.04$  and  $P = 0.001$ , respectively). Markers of NK and NKT cells, dendritic cells (immature and activated), and tumor-associated macrophages were also grouped in clusters 2 and 3 (Fig. 3A, bottom).

**Figure 4.** Disease-free survival of colorectal cancer patients according to expression of genes. A to C, disease-free survival of 103 patients according to high (red lines) or low (black lines) mRNA expression levels of IRF1 (A), GNLY (B), and VEGF (C) genes. D, disease-free survival of patients according to the expression levels of the GNLY gene in combination with IRF1 (top) and VEGF (bottom) genes (thin red lines, GNLY-Hi/IRF1-Hi, GNLY-Hi/VEGF-Hi; bold red lines, GNLY-Hi/IRF1-Lo, GNLY-Hi/VEGF-Lo; thin black lines, GNLY-Lo/IRF1-Hi, GNLY-Lo/VEGF-Hi; bold black lines, GNLY-Lo/IRF1-Lo, GNLY-Lo/VEGF-Lo). The cutoff value for the expression of each gene was defined at the median of the cohort. \*,  $P < 0.05$ , log-rank test.





**Figure 5.** Proposed model: control of colorectal cancer outcome by the immune system. Before surgery, immune strength and coordination are in balance with mechanisms of tumor escape (tumor immunogenicity, inflammation, and angiogenesis) to control metastatic invasion from the primary tumor site. Four major immune coordination profiles within colorectal cancer primary tumors are found: (a) Strong and coordinated adaptive immune responses mediated by cytotoxic (blue cells; *GZMY*) and Th1 (yellow cells; *IRF1*) effector memory T cells (green cells) may contribute to the elimination of migrating tumor cells (red cells). (b) Angiogenic (*VEGF*) and inflammatory processes may facilitate metastatic invasion and (c) noncoordinate immune responses. (d) Weak (*Lo*) immune reactions (immune ignorance?). After surgical removal of clinically detectable tumors, the parameters defining this balance are significantly associated with the risks of cancer relapse [2-y disease-free survival (*DFS*)]. It could be postulated that the amount of invading occult tumors and the amount of circulating memory T cells, generated within distinct primary tumor microenvironment, are in balance to control cancer re-emergence in the periphery (after surgery).

Compared with Hi patients, META- Lo (Fig. 3B, top) and META+ Lo (Fig. 3B, center and bottom) patients had very distinct correlation profiles with a majority of noncorrelated markers (yellow). Furthermore, except for the only significant positive correlation between final effector and CD8<sup>+</sup> T-cell subpopulations (CD8<sup>+</sup>/CD127<sup>-</sup>CD62L<sup>-</sup>;  $P = 0.03$ ) in the META+ Lo matrix, all patterns of significant positive correlations observed in Hi patients were lost in Lo patients (Supplementary Fig. S2).

This analytic approach allowed us to visualize the absence of immune coordination in patients with low intratumoral T-cell densities, whereas patients with high intratumoral T-cell densities presented correlation patterns consistent with continual recruitment and proliferation of activated CD8 T cells associated with complete memory T-cell differentiation at the primary tumor site (CD8 T-cell/effector memory T-cell correlations). This profile of efficient immune reaction is in balance (negative correlation) with patterns that could illustrate altered immune responses (CD4/early memory T-cell/B-cell correlations). Because in Hi patients the presence of metastases was associated with (a) a significant decrease of CD8 and late memory T cells and innate cells and (b) a significant increase of B cells (Figs. 1 and 2), our data suggest altered immune reactions in META+ Hi patients.

**Prognostic value of cellular immune coordination.** We next assessed the effect of immune coordination on the proliferation/apoptosis status of primary tumor cells by Ki67/M30 immunohistochemical stainings of cognate tumor samples (Supplementary Fig. S4). No differences were observed among the patient groups, suggesting that the effect on cancer progression of the immune system may be inefficient for the destruction of the primary tumor. To validate the effect of the coordination of *in situ* immune response on colorectal cancer prognosis, we evaluated the density of intratumoral immune T cells in a large cohort of 435 patients.

We investigated the CD8/CD3 T-cell density ratio in relation to clinical outcome in TMA experiments. Increased densities of T-cell infiltrates exhibiting high proportions of CD8 cytotoxic T cells within the primary tumor of colorectal cancer patients were associated with a significant protection against tumor recurrence (Supplementary Fig. S5).

To better characterize the mechanisms involved in antitumoral activity at the tumor-host interface, we investigated the effect on clinical outcome of mRNA expression levels of 17 mediators involved in immune or tumoral mechanisms. For each gene, patients were defined as high or low according to median gene expression. Disease-free survival rates were then calculated for each patient group. The prognostic value of the expression levels of genes related to T-cell populations (*CD3 $\zeta$* , *CD4*, *CD8 $\alpha$* ), Th1 adaptive immune responses (*TBX21/T-BET*, *IRF1*, *IFN $\gamma$* ), and cytotoxicity (*GZMY*, *GZMB*) were assessed. High expression of *TBX21/T-BET*, *IFN $\gamma$* , *IRF1*, and *GZMY* was associated with significantly improved disease-free survival rates ( $P = 0.02$ ,  $P = 0.02$ ,  $P = 0.0003$ , and  $P = 0.0004$ , respectively). Disease-free survival Kaplan-Meier curves according to *GZMY* and *IRF1* gene expression were illustrated (Fig. 4A and B, respectively). Conversely to immune mediators, the expressions of cancer-promoting genes involved in tumor invasion (*CEACAM1*), metastasis spreading (*EBAG9* and *CEA*), tumor cell antiapoptosis (*BIRC5/Survivin*), immune suppression (*IL-10* and *TGF $\beta$* ; data not shown), and angiogenesis (*VEGF*; Fig. 4C) had no prognostic values.

Immune coordination at the molecular level was assessed by analyzing combined expression of genes. We found significantly improved disease-free survival rates in patients with high combined gene expressions (Hi/Hi) of marker combinations related to CD4 T cells of Th1 phenotype (CD4/T-BET, CD4/IFN $\gamma$ , CD4/IRF1 patients) and cytotoxic CD8 T cells (CD8/GZMY)

compared with patients expressing low levels of these genes (Lo/Lo;  $P = 0.04$ ,  $P = 0.002$ ,  $P = 0.002$ , and  $P = 0.004$ , respectively; Supplementary Fig. S6). High coexpression of *IRF1* and *GNL1* genes (GNLY-Hi/IRF1-Hi) was essential for beneficial outcome with median disease-free survival >140 months, whereas patients expressing low levels of one of these genes or both (GNLY-Hi/IRF1-Lo, GNLY-Lo/IRF1-Hi, GNLY-Lo/IRF1-Lo) had median disease-free survival <15 months (Fig. 4D, top). These observations confirmed the importance of a strong coordination between immune mediators of cytotoxic and Th1 adaptive immune responses for favorable colorectal cancer outcome.

Finally, we assessed the effects of the tumor microenvironment on *in situ* antitumoral immune responses. We analyzed the prognostic value of the expression of protumoral mediators in combination with *GNLY*. Among all tested genes, only *VEGF* showed a profound effect on patient survival when expressed with *GNLY*. Patients expressing high levels of *GNLY* and *VEGF* genes (GNLY-Hi/VEGF-Hi) and patients expressing low levels of *GNLY* (GNLY-Lo/VEGF-Hi and GNLY-Lo/VEGF-Lo) had similar disease-free survival rates that were significantly lower than the disease-free survival rates of GNLY-Hi/VEGF-Lo patients ( $P < 0.004$  for all comparisons; Fig. 4D, bottom).

## Discussion

The mechanisms controlling tumor progression and cancer relapse are not clearly characterized. Here, we investigated the quality of the immune reaction at the primary tumor site during cancer progression (i.e., according to the density of tumor-infiltrating T cells and the metastatic status of the patients). We showed that coordination of the immune response was drastically impaired in patients with low densities of intratumoral T cells compared with patients with high densities of such cells. Phenotypic correlation analyses showed matrices that were highly fragmented with no particular functional relevance in both META-Lo and META+ Lo patients. This suggested the absence of coordinated immune response independently of the metastatic status in Lo patients. Conversely, a high density of tumor-infiltrating T cells (Hi patients) was associated with strong immune coordination. Significant positive correlations between T cells of late memory and cytotoxic phenotypes indicated continual recruitment, activation, and memory differentiation of CD8 T cells at the primary tumor site. In larger cohorts of patients using tissue microarrays, we also showed that high densities of T cells associated with a high CD8/CD3 density ratio correlated with a very good prognosis. In contrast, low adaptive immune coordination was associated with very poor prognosis. Consistently, patients presenting high and coordinated intratumoral expression of the global Th1 immune response marker *IRF1* and the cytotoxicity-specific marker *GNLY* had significantly better survival rates compared with patients expressing heterogeneous or low levels of these genes.

Yet, in Hi patients, the presence of metastases was associated with (a) a significant decrease of innate immune cells, (b) a significant decrease of CD8 T cells and fully differentiated memory T cells, (c) loss of the phenocluster of markers of innate cells and early activated T cells (illustrating innate/adaptive immune compartment interactions), and (d) a significant increase of B cells (suggesting immune deviation mechanisms; refs. 34, 35).

According to the coexpression of *IRF1* and *GNLY*, the frequencies of strong immune coordination parameters were

reduced in META+ patients (data not shown). META+ patients represented only 48% of GNLY-Hi/IRF1-Hi patients and 65% of GNLY-Lo/IRF1-Lo patients. Overall, these observations represent clues of altered immune responses when metastases are present. However, a significant number of META+ patients displayed a high degree of immune coordination preventing relapse events. This raises two hypotheses: Does the alteration of the immune reaction at the primary tumor site facilitate metastatic invasion? Is the immune system overwhelmed and affected by the presence of metastases? Interestingly, some patients without lymph node and/or distant organs (META- patients) have an absence of immune coordination and low densities of T cells. Thus, local immune escape mechanisms may exist in the primary tumor even before metastatic spread. Because proliferation and apoptosis rates of tumor cells were not significantly different between the patient groups, the outgrowth of the primary tumor may overcome the destruction by the immune system. Whichever hypothesis on the mechanisms of long-term relapse prevention after surgery may be related to the quality of the immune reaction at the primary tumor site even in patients with advanced colorectal cancer.

Among the distinct factors potentially involved in immune escape at the primary site, several mechanisms or cell types may participate, such as immature dendritic cells, regulatory T cells, Th1/Th2 immune response switch, immunosuppression, local metastatic invasion, inflammation, and angiogenesis. In patients with metastases or low intratumoral T-cell densities, there was no increased expression of markers of regulatory T cells, tumor-associated macrophages, and immature dendritic cells by flow cytometry (CD4<sup>+</sup>CD25<sup>hi</sup>, CD45<sup>+</sup>CD14<sup>+</sup>, and CD45<sup>+</sup>CD1a<sup>+</sup>CD14<sup>-</sup>CD83<sup>-</sup>, respectively) and by tissue microarray (Foxp3<sup>+</sup>, CD68<sup>+</sup>, and CD1a<sup>+</sup>, respectively) experiments (data shown). This may suggest the existence of immune ignorance or reduced tumor immunogenicity mechanisms for differential immune cell recruitment among patients. Interestingly, we found that only the proangiogenic factor VEGF had a deleterious effect on relapse prevention mechanisms associated with strong antitumoral immune reaction. *VEGF* expression levels had no prognostic value per se, in agreement with immunohistochemical-based studies (36, 37). Our data indicate that cytotoxic Th1 adaptive immune responses may be necessary, but not sufficient, to prevent tumor recurrence. At the primary tumor site, inflammatory cytokines (such as IL-1A, IL-6, IL-8, oncostatin M, and tumor necrosis factor  $\alpha$ ) can enhance tumorigenic processes by up-regulating important mediators of angiogenesis, such as VEGF (38). In this context, the effect on colorectal cancer outcome of the balance between *GNLY/IRF1* and *VEGF* expressions may reflect beneficial cytotoxic Th1 adaptive immune responses versus deleterious inflammatory reaction (39–41). However, other roles of angiogenesis may affect cytotoxic Th1 adaptive immune responses. In the primary tumor site, the role of angiogenesis in promoting nutrient supply (21) may not explain the obliteration of the beneficial role of strong immune responses. In contrast, the induction of vascular exit paths for migrating tumor cells (42) could result in increased metastatic dissemination favoring relapse occurrence. In this case, even strong *in situ* cytotoxic Th1 adaptive immune responses may not be sufficient to counteract metastatic invasion. Thus, in the periphery, great number of occult tumor cells may overwhelm immunosurveillance mechanisms during the equilibrium phase. Furthermore, if the migrating tumor cells inherit the strong angiogenic properties of their resident counterparts, occult tumor outgrowth may be further enhanced (43). This idea that angiogenesis

and adaptive immune responses are strongly linked in cancer recurrence should be taken into account when considering therapeutic options.

We were able to describe four major immune coordination profiles within colorectal cancer primary tumors depending on the balance between tumor escape and immune coordination: (a) strong and coordinate cytotoxic Th1 immune responses (*GNL1/IRF1*) without or (b) with tumor angiogenesis (*VEGF*), (c) noncoordinate immune responses, and (d) weak (Lo) immune reactions (immune ignorance?). These distinct immune profiles are associated with significant distinct cancer outcome (relapse risks), as summarized in Fig. 5.

It is suspected that metastatic invasion can lead to the dissemination of tumor cells that can remain in an asymptomatic and nondetectable state of dormancy (i.e., not expanding in mass) for long periods of time before cancer re-emergence (44). Control of cancer dormancy involves various mechanisms such as cellular dormancy ( $G_0$ - $G_1$  arrest), angiogenic dormancy, and immunosurveillance (45). Recently, Koebel and colleagues (46) showed that stable lesions of transformed immunogenic cells in mice were controlled by the adaptive immune system of the host in a condition of "equilibrium." In these experiments, loss of either immunocompetence or immunogenicity could lead to tumor outgrowth. We previously showed that the absence of microscopic evidence of early metastatic invasiveness within lymphovascular vessels was associated with high densities of effector memory T cells within primary tumors and that both criteria were powerful indicators of improved prognosis in human colorectal cancer (17). Based on these data, it could be proposed that the immune system exerts its protective role against cancer relapse (a) at the primary tumor site by eliminating migrating tumor cells, subsequently reducing the number of disseminated occult tumors, and (b) in the periphery by controlling occult tumor evolution from dormancy state to cancer re-emergence (equilibrium phase). Moreover, these two antitumoral functions of the immune system could be tightly associated. As suggested in mice (47), cytotoxic effector memory T cells reacting at the primary tumor site might also, after surgical

removal of tumors, be in charge of long-term antitumoral immunity in colorectal cancer.

In conclusion, our study argues for the involvement of immune coordination and late memory and cytotoxic T-cell populations in antitumoral activity against human colorectal cancer (Fig. 5). First, due to their enhanced cytotoxic capabilities, effector memory T cells may be involved in the control of metastatic invasion at the primary tumor site. Second, due to their memory properties, effector memory T cells may provide long-term protection against outgrowth of disseminated occult tumor cells potentially involved in relapse events. Depending on the strength and coordination of antitumoral immune responses elicited in primary tumor microenvironments (level of immunogenicity and angiogenesis), populations of T cells with distinct quantity (number of clones) and quality (memory differentiation state) could be generated. Subsequently, distinct potentials for long-lived antitumoral immunity may be maintained after surgical resection of primary and secondary tumors. Future comparative studies of tumors according to immune parameters and angiogenesis may reveal biological mechanisms involved in emergence and cancer progression. More adapted treatment and therapeutic strategies may ultimately be proposed to cure colorectal cancer.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Acknowledgments

Received 7/10/2008; revised 12/11/2008; accepted 1/2/2009; published OnlineFirst 3/3/09.

**Grant support:** Association pour la Recherche sur le Cancer (ARC), the National Cancer Institute (INCa), the Canceropole Ile de France, Ville de Paris, Immucan, INSERM, the Austrian Federal Ministry of Science and Research (GEN-AU project Bioinformatics Integration Network) and Academic Cooperation and Mobility Unit (fellowship to P. Charoentong), and the European Commission (7FP, Geninca Consortium, grant no. 202230).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

## References

- Finn OJ. Cancer immunology. *N Engl J Med* 2008;358:2704-15.
- Dighe AS, Richards E, Old LJ, Schreiber RD. Enhanced *in vivo* growth and resistance to rejection of tumor cells expressing dominant negative IFN $\gamma$  receptors. *Immunity* 1994;1:447-56.
- Kaplan DH, Shankaran V, Dighe AS, et al. Demonstration of an interferon  $\gamma$ -dependent tumor surveillance system in immunocompetent mice. *Proc Natl Acad Sci U S A* 1998;95:7556-61.
- Smyth MJ, Thia KY, Street SE, MacGregor D, Godfrey DI, Trapani JA. Perforin-mediated cytotoxicity is critical for surveillance of spontaneous lymphoma. *J Exp Med* 2000;192:755-60.
- van den Broek ME, Kagi D, Ossendorp F, et al. Decreased tumor surveillance in perforin-deficient mice. *J Exp Med* 1996;184:1781-90.
- Shankaran V, Ikeda H, Bruce AT, et al. IFN $\gamma$  and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* 2001;410:1107-11.
- Zhou G, Lu Z, McCadden JD, Levitsky HI, Marson AL. Reciprocal changes in tumor antigenicity and antigen-specific T cell function during tumor progression. *J Exp Med* 2004;200:1581-92.
- Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoeediting. *Annu Rev Immunol* 2004;22:329-60.
- Clemente CG, Mihm MC, Jr., Bufalino R, Zurrada S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. *Cancer* 1996;77:1303-10.
- Sato E, Olson SH, Ahn J, et al. Intraepithelial CD8 $^+$  tumor-infiltrating lymphocytes and a high CD8 $^+$ /regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proc Natl Acad Sci U S A* 2005;102:18538-43.
- Zhang L, Conejo-Garcia JR, Katsaros D, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* 2003;348:203-13.
- Baier PK, Wimmenauer S, Hirsch T, et al. Analysis of the T cell receptor variability of tumor-infiltrating lymphocytes in colorectal carcinomas. *Tumour Biol* 1998;19:205-12.
- Dalerba P, Maccalli C, Casati C, Castelli C, Parmiani G. Immunology and immunotherapy of colorectal cancer. *Crit Rev Oncol Hematol* 2003;46:33-57.
- Diederichsen AC, Hjelmberg JB, Christensen PB, Zeuthen J, Fenger C. Prognostic value of the CD4 $^+$ /CD8 $^+$  ratio of tumour infiltrating lymphocytes in colorectal cancer and HLA-DR expression on tumour cells. *Cancer Immunol Immunother* 2003;52:423-8.
- Naito Y, Saito K, Shiiba K, et al. CD8 $^+$  T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Res* 1998;58:3491-4.
- Prall F, Duhrop T, Weirich V, et al. Prognostic role of CD8 $^+$  tumor-infiltrating lymphocytes in stage III colorectal cancer with and without microsatellite instability. *Hum Pathol* 2004;35:808-16.
- Pages F, Berger A, Camus M, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* 2005;353:2654-66.
- Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006;313:1960-4.
- Galon J, Fridman WH, Pages F. The adaptive immunologic microenvironment in colorectal cancer: a novel perspective. *Cancer Res* 2007;67:1883-6.
- Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57-70.
- Bergers G, Benjamin LE. Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* 2003;3:401-10.
- Smyth MJ, Godfrey DI, Trapani JA. A fresh look at tumor immunosurveillance and immunotherapy. *Nat Immunol* 2001;2:293-9.
- Dunn GP, Koebel CM, Schreiber RD. Interferons, immunity and cancer immunoeediting. *Nat Rev Immunol* 2006;6:836-48.
- Kim R, Emi M, Tanabe K, Arihiro K. Tumor-driven evolution of immunosuppressive networks during malignant progression. *Cancer Res* 2006;66:5527-36.
- Zou W. Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat Rev Cancer* 2005;5:263-74.



26. Khong HT, Restifo NP. Natural selection of tumor variants in the generation of "tumor escape" phenotypes. *Nat Immunol* 2002;3:999-1005.
27. Mocellin S, Wang E, Marincola FM. Cytokines and immune response in the tumor microenvironment. *J Immunother* 2001;24:392-407.
28. Egen JG, Kuhns MS, Allison JP. CTLA-4: new insights into its biological function and use in tumor immunotherapy. *Nat Immunol* 2002;3:611-8.
29. Okazaki T, Honjo T. The PD-1-PD-L pathway in immunological tolerance. *Trends Immunol* 2006;27:195-201.
30. Sobin LH, Greene FL. Global TNM advisory group. *Cancer* 2004;100:1106.
31. Boulton SJ, Gartner A, Reboul J, et al. Combined functional genomic maps of the *C. elegans* DNA damage response. *Science* 2002;295:127-31.
32. Galon J, Franchimont D, Hiroi N, et al. Gene profiling reveals unknown enhancing and suppressive actions of glucocorticoids on immune cells. *FASEB J* 2002;16:61-71.
33. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics* 2002;18:207-8.
34. Shah S, Divekar AA, Hilchey SP, et al. Increased rejection of primary tumors in mice lacking B cells: inhibition of anti-tumor CTL and TH1 cytokine responses by B cells. *Int J Cancer* 2005;117:574-86.
35. Tan TT, Coussens LM. Humoral immunity, inflammation and cancer. *Curr Opin Immunol* 2007;19:209-16.
36. Doger FK, Meteoglu I, Tuncyurek P, Okyay P, Cevikel H. Does the EGFR and VEGF expression predict the prognosis in colon cancer? *Eur Surg Res* 2006;38:540-4.
37. Zheng S, Han MY, Xiao ZX, Peng JP, Dong Q. Clinical significance of vascular endothelial growth factor expression and neovascularization in colorectal carcinoma. *World J Gastroenterol* 2003;9:1227-30.
38. Angelo LS, Kurzrock R. Vascular endothelial growth factor and its relationship to inflammatory mediators. *Clin Cancer Res* 2007;13:2825-30.
39. Balkwill F, Coussens LM. Cancer: an inflammatory link. *Nature* 2004;431:405-6.
40. Moore RJ, Owens DM, Stamp G, et al. Mice deficient in tumor necrosis factor- $\alpha$  are resistant to skin carcinogenesis. *Nat Med* 1999;5:828-31.
41. Voronov E, Shouval DS, Krelin Y, et al. IL-1 is required for tumor invasiveness and angiogenesis. *Proc Natl Acad Sci U S A* 2003;100:2645-50.
42. Folkman J. Role of angiogenesis in tumor growth and metastasis. *Semin Oncol* 2002;29:15-8.
43. Naumov GN, Akslen LA, Folkman J. Role of angiogenesis in human tumor dormancy: animal models of the angiogenic switch. *Cell Cycle* 2006;5:1779-87.
44. Vessella RL, Pantel K, Mohla S. Tumor cell dormancy: an NCI workshop report. *Cancer Biol Ther* 2007;6:1496-504.
45. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer* 2007;7:834-46.
46. Koebel CM, Vermi W, Swann JB, et al. Adaptive immunity maintains occult cancer in an equilibrium state. *Nature* 2007;450:903-7.
47. Xiang R, Lode HN, Gillies SD, Reisfeld RA. T cell memory against colon carcinoma is long-lived in the absence of antigen. *J Immunol* 1999;163:3676-83.

## Article 3

# "Biomolecular network reconstruction identifies T cell homing factors associated with survival in colorectal cancer"

Bernhard Mlecnik, Marie Tosolini, Pornpimol Charoentong, Amos Kirilovsky, Gabriela Bindea, Anne Berger, Matthieu Camus, Mélanie Gillard, Patrick Bruneval, Wolf-Herman Fridman, Franck Pagès, Zlatko Trajanoski, Jérôme Galon

Gastroenterology, *in press*

Le cancer colorectal est une maladie complexe impliquant des mécanismes aussi divers que l'accumulation d'altérations génétiques ou l'interaction tumeur-mécanismes de défense de l'hôte. Le système immunitaire est lui aussi complexe. Or les systèmes immunitaires innés et adaptatifs peuvent protéger l'hôte contre le développement de la tumeur. Toute cette complexité doit être prise en compte. Pour cela, nous utilisons une stratégie qui permet l'intégration des données, une reconstruction des réseaux moléculaires et une création d'hypothèses ciblant la détection de réseaux immunitaires importants pour le pronostic des patients.

J'ai effectué les marquages des lymphocytes Th1 (T-Bet), des macrophages (CD68) et des cellules dendritiques immatures (CD1a) par immunohistochimie sur les Tissue MicroArrays de 415 patients. J'ai analysé ces différents marquages à l'aide du logiciel Spotbrowser au centre et front de progression de la tumeur (figure 3D, tableau 1, figure supplémentaire 3 et tableau supplémentaire 4). J'ai également effectué les extractions d'ARN puis les PCR en temps réel sur Low-Density-Array notamment pour l'expression des chimiokines CX3CL1, CXCL10 et CXCL9 et les molécules d'adhésion ICAM1, VCAM1 et MADCAM1 (figure 2, 3 et 5).

## Résumé

Le cancer colorectal est une maladie complexe impliquant des mécanismes de défense immunitaire contre la tumeur. Nous avons précédemment montré l'importance de la réponse immunitaire et notamment d'une forte densité de lymphocytes T (CD3+) et T mémoire (CD45RO+) (article 1 [140]). Nous avons utilisé l'intégration de données et la reconstruction de réseaux moléculaires pour créer des hypothèses sur les mécanismes sous-jacents les réponses immunitaires relevantes pour le pronostic des patients. Ces hypothèses ont été ensuite vérifiées.

Cette approche nous a révélé l'importance de trois chimiokines (CX3CL1, CXCL10 et CXCL9) ainsi que de trois molécules d'adhésions (ICAM1, VCAM1 et MADCAM1) dans l'attraction et donc la présence d'une forte densité de cellules immunitaires. Une forte expression de ces molécules corrèle avec une survie sans récurrence prolongée. De plus, une forte expression des chimiokines CX3CL1, CXCL10 et CXCL9 corrèle avec une forte densité de lymphocytes T CD8+ ( $p \leq 0,05$ ). De plus, ces chimiokines corrèlent avec des sous-populations lymphocytaires particulières. En effet, CXCL10 et CXCL9, contrairement à CX3CL1, corrèlent avec la présence de lymphocytes T CD8 mémoires (CD3+ CD8+ CD45RO+). Les tumeurs de patients ayant une expression élevée de CX3CL1 ont une densité significativement plus élevée de cellules cytotoxiques ( $P \leq 0,05$  pour granzyme B et CD57 au centre et au front de progression de la tumeur). CX3CL1 semble attirer également une sous-population T avec des TCR particuliers. En effet, en présence d'une forte expression de CX3CL1, les niveaux des récepteurs  $V\beta L03$ ,  $V\beta L07$ ,  $V\beta L06$  et  $V\beta L09$  sont augmentés ( $P \leq 0,01$ ,  $P \leq 0,008$ ,  $P \leq 0,05$  et  $P \leq 0,05$ , respectivement). Ces TCR particuliers corrèlent avec le bon pronostic des patients.

La stratégie permettant l'intégration des données, une reconstruction des réseaux moléculaires et une création d'hypothèses est une stratégie puissante pour découvrir les mécanismes moléculaires sous-jacents. Cela nous a permis de démontrer l'importance de chimiokines et molécules d'adhésion pour la présence de fortes densités de lymphocytes T au site de la tumeur.

## Accepted Manuscript

Biomolecular network reconstruction identifies T cell homing factors associated with survival in colorectal cancer

Bernhard Mlecnik, Marie Tosolini, Pornpimol Charoentong, Amos Kirilovsky, Gabriela Bindea, Anne Berger, Matthieu Camus, Mélanie Gillard, Patrick Bruneval, Wolf-Herman Fridman, Franck Pagès, Zlatko Trajanoski, Jérôme Galon



PII: S0016-5085(09)01960-X  
DOI: 10.1053/j.gastro.2009.10.057  
Reference: YGAST 55846

To appear in: *Gastroenterology*

Received date: 4 August 2009  
Revised date: 12 October 2009  
Accepted date: 29 October 2009

Please cite this article as: Mlecnik, B., Tosolini, M., Charoentong, P., Kirilovsky, A., Bindea, G., Berger, A., Camus, M., Gillard, M., Bruneval, P., Fridman, W.H., Pagès, F., Trajanoski, Z., Galon, J., Biomolecular network reconstruction identifies T cell homing factors associated with survival in colorectal cancer, *Gastroenterology* (2009), doi: 10.1053/j.gastro.2009.10.057.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Biomolecular network reconstruction identifies T cell homing factors  
associated with survival in colorectal cancer**

Short title: Biomolecular networks and tumor-host interactions

Bernhard Mlecnik,<sup>1,2,3,4</sup> Marie Tosolini,<sup>1,2,3</sup> Pornpimol Charoentong,<sup>4</sup> Amos Kirilovsky,<sup>1,2,3</sup>  
Gabriela Bindea,<sup>1,2,3,4</sup> Anne Berger,<sup>5</sup> Matthieu Camus,<sup>1,2,3</sup> Mélanie Gillard,<sup>1,2,3</sup> Patrick  
Bruneval,<sup>6</sup> Wolf-Herman Fridman,<sup>1,2,3,7</sup> Franck Pagès,<sup>1,2,3,7</sup> Zlatko Trajanoski,<sup>4,\*</sup> Jérôme  
Galon,<sup>1,2,3,5,\*</sup>

<sup>1</sup> INSERM, Integrative Cancer Immunology Team, INSERM U872, 75006 Paris, France;

<sup>2</sup> Université Paris Descartes, Paris, France; <sup>3</sup> Cordeliers Research Center, Université  
Pierre et Marie Curie Paris 6, Paris, France. <sup>4</sup> Institute for Genomics and Bioinformatics,  
Graz University of Technology, Graz, Austria. Departments of <sup>5</sup> General and Digestive  
Surgery, and <sup>6</sup> Pathology, and <sup>7</sup> Immunology, Georges Pompidou European Hospital,  
Assistance Publique-Hôpitaux de Paris, AP-HP, Paris, France.

\* Correspondence: jerome.galon@crc.jussieu.fr and zlatko.trajanoski@tugraz.at

There are no conflict of interest to disclose for all authors

list of how each author was involved with the manuscript:

Bernhard Mlecnik (acquisition of data; analysis and interpretation of data; drafting and revision of the manuscript; statistical analysis)

Marie Tosolini (acquisition of data; interpretation of data; technical; revision of the manuscript)

Pornpimol Charoentong (analysis and interpretation of data)

Amos Kirilovsky (acquisition of data; interpretation of data; technical)

Gabriela Bindea (interpretation of data; statistical analysis)

Anne Berger (material support)

Matthieu Camus (acquisition of data; technical)

Mélanie Gillard (technical)

Patrick Bruneval (material support)

Wolf-Herman Fridman (interpretation of data; critical revision of the manuscript)

Franck Pagès (analysis and interpretation of data; critical revision of the manuscript)

Zlatko Trajanoski (interpretation of data; critical revision of the manuscript for important intellectual content; obtained funding)

Jérôme Galon (study concept and design; analysis and interpretation of data; writing of the manuscript; study supervision; obtained funding)

Title: 114 characters; Abstract: 221 words; Text: 5981 words

**Abstract**

**BACKGROUND & AIMS:** Colorectal cancer is a complex disease involving immune defense mechanisms within the tumor. Herein, we used data integration and biomolecular network reconstruction to generate hypotheses about the mechanisms underlying immune responses in colorectal cancer that are relevant to tumor recurrence.

**METHODS:** Mechanistic hypotheses were formulated on the basis of data from 108 patients and tested using different assays (gene expression, phenome mapping, tissue-microarrays, TCR-repertoire).

**RESULTS:** This integrative approach revealed chemoattraction and adhesion to play important roles in determining the density of intratumoral immune cells. The presence of specific chemokines (CX3CL1, CXCL10, CXCL9) and adhesion molecules (ICAM1, VCAM1, MADCAM1) correlated with different subsets of immune cells and with high-densities of T-cell subpopulation within specific tumor regions. High expression of these molecules correlated with prolonged disease-free survival. Moreover, the expression of certain chemokines associated with particular TCR-repertoire and specific TCR usage predicted patient survival.

**CONCLUSIONS:** Data integration and biomolecular network reconstruction is a powerful approach to uncover molecular mechanisms. This study, demonstrates the utility of this approach for the investigation of malignant tumors and other diseases. In colorectal cancer, the expression of specific chemokines and adhesion molecules were found as being critical for high densities of T cell subsets within the tumor, and associated with particular TCR-repertoire. Intratumoral specific TCR usage correlated with the prognosis of the patients.

## Introduction

In order to develop stratified or personalized strategies for complex multifactorial diseases it is important to understand how numerous and diverse elements function together in human pathology<sup>1, 2</sup>. A comprehensive understanding of diseases such as cancer will not only require the integration and analysis of data from tumor in its microenvironment, but also of other data sources from model organisms stored in public databases<sup>1, 2</sup>. Cancer is the result of an accumulation of genetic alterations that allows growth of neoplastic cells<sup>3, 4</sup>. The adenoma-carcinoma sequence underlies the development of colorectal cancer (CRC), and distinct pathways (microsatellite instability (MSI) and chromosomal instability (CIN) pathways) have been identified<sup>5</sup>. The natural evolution of a cancer also involves antagonistic interactions of the tumor with the defense mechanisms of the host<sup>6, 7</sup>. Inflammatory mediators can promote tumor progression and metastases<sup>8</sup>. The innate and adaptive immune systems can also protect the host against tumor development through mechanisms of immunosurveillance<sup>9</sup>. The increased susceptibility of immunodeficient mice to carcinogen-induced and spontaneous tumors demonstrated the role of innate and adaptive immunity in the control of tumor development<sup>9-11</sup>. More recent data provide support for a role for adaptive immunity also during the equilibrium phase of cancer<sup>12</sup>.

In human CRC, adaptive immune reaction was found and densities of immune cells are very different from patient to patient<sup>7</sup>. Numerous HLA-restricted T-cells specific for tumor peptides have been described<sup>13</sup>. Lymphocytes infiltrating solid tumors have been associated with improved prognosis<sup>14-16</sup>. Tumors from CRC patients containing a high density of infiltrating memory and effector memory T-cells ( $T_{EM}$ ) were found to be less likely to disseminate to lymphovascular and perineural structures and to regional lymph-



nodes<sup>17</sup>. Tumor recurrence and overall patient survival times correlated broadly with the immune context and the presence of memory T cells within the tumor<sup>18, 19</sup>. Tumors also contain a variety of cytokines, chemokines and inflammatory and cytotoxic mediators. This complex network reflects the heterogeneity underlying tumor biology and tumor-host interactions<sup>7, 9</sup>. The reasons for the very different densities of immune cells found within tumors, however, remain unknown.

To gain an improved understanding of tumor host-interactions in human CRC, we have developed and applied an intuitive data integration strategy to analyze immune reaction in CRC. We used a method that effectively created hypotheses permitting us to detect immune network relevant to prognosis. Predicted molecules involved in lymphocyte chemoattraction and adhesion were analyzed together with immune populations *in situ*. Biological hypotheses were then validated in a large cohort of patients by a combination of high-throughput approaches. The novel aspects of our study revealed mechanisms resulting in high or low densities of specific immune cells at the tumor site. Chemokines and adhesion molecules associated with immune effector T cells with particular TCR-repertoire. Furthermore, the presence of a specific intratumoral TCR-repertoire correlated with the survival of the patient. Thus, we provided a framework for predicting effective host-immune reaction against cancer in man. This study, demonstrates the utility of data integration and biomolecular network reconstruction for the investigation of malignant tumors and other diseases.

## Results

### **Immune-related genes are associated with the absence of tumor recurrence**

We first investigated gene expression in colorectal tumors. We determined the median cut-off values for each gene, and performed survival analysis for up to 10 years after primary tumor resection. Logrank P-values associated with disease-free survival were then calculated and hazard ratios were illustrated by the size of each node in a network (Figure 1). The expression of genes associated with tumor invasion (CEACAM1, CD97), metastasis spreading (ACE, EBAG9, MMP7), tumor anti-apoptotic (survivin/BIRC5), and angiogenesis (VEGF) was assessed. Surprisingly, the duration of disease free survival (DFS) did not correlate significantly with the expression of these tumor-related genes. Host-immune response-related genes, particularly pro-inflammatory, immunosuppressive, T-helper-specific ( $T_H1$ ,  $T_H2$ ), innate and adaptive immune response-related genes were also assessed. The patterns of expression of pro-inflammatory (PTGS2, IRAK4),  $T_H2$  (GATA3) and immunosuppression (IL10, FoxP3) related genes did not vary according to tumor recurrence. In contrast, innate and adaptive immunity related-genes (GNLY, STAT1, IRF1, IFNG, TBX21, IL18RAP, ICOS),  $T_H1$  (STAT1, IRF1, IFNG, TBX21), as well as genes involved in T cell activation,  $T_H1$ , and negative regulation of the immune response (PDCD1, PDCD1LG1, PDCD1LG2) stratified patients into groups with statistically different DFS ( $P < 0.05$ ).

### **Reconstructed biomolecular network predicts interacting chemokines and adhesion molecules**

Based on the gene expression data we reconstructed a gene-gene network (see Supplementary Methods). Using the subset of genes relevant to tumor recurrence and with statistically different DFS (Supplementary Table 1), we further combined publicly

available databases and prior knowledge<sup>20</sup> to enrich the network. The prediction of genes was based on conserved genomic neighborhood, phylogenetic profiling, co-expression analysis, protein-protein interaction, functional genomic public databases, and literature co-occurrence. The reconstructed network was visualized (Figure 1) using a network layout visualization which uses Gene Ontology (GO) annotations as a source of external class information to direct the network layout process and to emphasize the biological function of the nodes (Supplementary Table 2)<sup>21</sup>.

This integration and visualization of both experimental and *in silico* data on the network revealed putative functional interactions and new groups of genes associated with the patient's prognosis. Among the prediction of network membership (nodes with red border) were molecules involved in leukocyte and myeloid cell differentiation, the regulation of apoptosis, the protein kinases cascade, adhesion and in chemotaxis (Supplementary Figure 1)<sup>22</sup>. To test whether the association between predicted genes and patient survival might be a result of their correlation with the seed genes (used for prediction), we performed STRING analysis without co-expression data. The network constructed without using co-expression information was highly similar to the initial network prediction (Supplementary Table 3).

The first top-ranked predicted gene was CX3CL1. Other chemokines, such as CXCL9, CXCL10, CCL2, CCL5, CCL11, and adhesion molecules such as MADCAM1, ICAM1 and VCAM1 were predicted to be interacting molecules (Figure 1).

### **Chemoattractants and adhesion molecules are associated with improved prognosis**

This reconstructed biomolecular network both generated testable hypotheses and predicted novel interactions. Among the predictions of network memberships were

chemokines. These molecules could attract distinct cell subpopulations associated with patient survival. To validate the predictions, we analyzed the gene expression of CX3CL1, CXCL9 and CXCL10 in primary tumors in two independent cohorts (n=108 and n=27). In the first cohort, an association between high chemokine expression and improved patient survival (cut-off at median of the dataset HR= 2.06, 1.78, and 1.76, respectively,  $P<0.05$ ) was observed for each marker (CX3CL1, CXCL9, and CXCL10). The hazard ratios (HR) for CX3CL1, CXCL9, and CXCL10 were increased (2.21, 2.38 and 2.92 respectively) by using the cut-off that yielded the minimum  $P$ -value for disease-free survival (Figure 2A). Similar results were also found in the second cohort (Supplementary Figure 2).

Patients with increased expression of other predicted chemokines and adhesion molecules such as CCL2, CCL5, CCL11, ICAM1 and MADCAM1 exhibited prolonged disease-free survival (HR Lo vs Hi from 1.81 to 2.21). In contrast, patients with decreased expression of EGFR showed prolonged disease-free survival. For control purposes, we tested the expression of a non-predicted chemokine, CXCL5. The expression of this chemokine did not vary according to tumor recurrence (Figure 2A). In order to investigate whether the combined analysis of predicted genes could improve the prediction of patient prognosis, we plotted a two-dimensional hierarchical cluster matrix according to Hazard Ratio between patient groups (Figure 2B). Using this method, the disease-free survival time differences found between patients (HiHi vs LoLo), were larger than those found by single chemokine analysis (Hi vs Lo) (HR= 2.30 and 1.78, respectively). The combined analysis of predicted adhesion molecules and chemokines revealed a statistically different disease-free survival as illustrated for ICAM1/CX3CL1, ICAM1/CXCL10 and CXCL9/MADCAM1 (Figure 2, A and B). Combined analysis

revealed four major clusters. High levels of genes from cluster A or a combination of genes from cluster A had little impact on patient survival. Patients with low levels of EGFR or high levels of CXCL10, CX3CL1, CCL11, MADCAM1, or CXCL9, (cluster B) had a very favorable prognosis compared to the whole cohort (all HR<0.61). High levels of the chemokine CX3CL1, together with high levels of CXCL9, CXCL10, CCL2, CCL5, CCL11, VCAM1, ICAM1 or MADCAM1 further increased disease-free survival (HR ranging from 0.54 to 0.29). In contrast, patients with a combination of genes from clusters C and D had a poor prognosis compared to the whole cohort (HR from 1.33 to 2.37). Thus, Kaplan-Meier curves and hazard ratio matrix displaying the duration of disease-free survival according to the gene combinations showed that an improved prognosis is associated with the expression of specific chemoattractants and adhesion molecules.

#### **Phenotypes of intratumoral immune cells correlated with disease-free survival**

The intratumoral immune cell infiltrate varies greatly between patients with colorectal cancer. In order to understand the reasons for this heterogeneity, the distribution of infiltrating immune cells was determined in tumors with high and low chemokine gene expression levels (CX3CL1, CXCL9, CXCL10). Immunostaining of tissue microarrays for CD8 T cell effectors was performed as illustrated in Figure 3 for two representative patients with respectively high and low levels of chemokine gene expression. The tumors of patients with high expression levels of CX3CL1, CXCL9 and CXCL10 were found to contain a significantly higher density of CD8 T cells ( $P<0.05$ ) (Figure 3A). To further investigate functional patterns and the coordination of immune cell populations within the primary tumor in relation to CX3CL1, CXCL9 and CXCL10 gene expression, we performed a phenotypic analysis of cell surface markers of the same tumors by flow

cytometry. Increased memory CD3+CD45RO+ T cell infiltration was observed for patients with high expression levels of CXCL9 and CXCL10, but not of CX3CL1. Increased CD3+CD8+ and CD3+TCR $\alpha\beta$ + T cell infiltration was furthermore detected for patients with high CX3CL1, CXCL9 and CXCL10 expression levels (Figure 3B). Pairwise comparisons were performed by measuring the similarity between profiles using Pearson correlation coefficients. The results of the correlation analysis were visualized using hierarchical clustering of a correlation matrix<sup>18</sup> of 149 cell surface markers analyzed by flow cytometry (Figure 3C). The correlation matrix revealed 6 major clusters. CX3CL1 exhibited positive correlations (cluster D,  $P < 0.05$  for all combinations) with the total density of infiltrating T cells (Total T cells, Total CD3+CD8+, Total CD3+CD4+). CXCL9 and CXCL10 showed a strong positive correlation (cluster C,  $R = 0.76$ ,  $P < 0.01$ ) and clustered together with a subpopulation of memory CD8 T cells (CD3+CD8+CD45RO+). Other immune cell populations were located in different clusters. Naïve T cells (Total CD3+CD45RA+CD27+) for instance were located in cluster A and NK (Total CD3-CD56+) and B cells (Total CD19+) in cluster F. These results thus indicate a high degree of functional coordination of specific types of intra-tumoral immune cells with the expression levels of specific subsets of chemokines.

To confirm these results, we investigated the density of immune cell populations in two specific regions of colorectal tumors using Tissue microarrays: the center (CT) and the invasive margin (IM) of the tumor. Total T lymphocytes (CD3), CD8 T cell effectors and their associated cytotoxic molecule (GZMB), subset of activated cytotoxic T cells and NK cells (CD57), memory T cells (CD45RO), T<sub>H</sub>1 cells (T-Bet), immature dendritic cells (CD1A), and macrophages (CD68), were in each case quantified by

immunostaining (Figure 3D). A significant correlation was observed between specific immune cell densities and chemokine expression levels (Table 1 and Supplementary Figure 3 and Supplementary Table 4). The tumors of patients with high CX3CL1 expression levels contained a significantly increased density of effector-activated cytotoxic T cells and T<sub>H</sub>1 cells *in situ* ( $P < 0.05$  for GZMB<sub>CT</sub>, T-Bet<sub>CT</sub>, CD8<sub>CT</sub>, CD57<sub>CT</sub>, CD3<sub>CT</sub>, and for GZMB<sub>IM</sub>, T-Bet<sub>IM</sub>, CD57<sub>IM</sub>). The tumors of patients with high CXCL9 and CXCL10 expression levels, in contrast, contained a significantly increased number of memory CD8 T cells and macrophages *in situ* ( $P < 0.05$  for CXCL9 and CD8<sub>CT</sub>, CD45RO<sub>CT</sub>, CD68<sub>CT</sub>, CD8<sub>IM</sub>, CD45RO<sub>IM</sub>, and  $P < 0.05$  for CXCL10 and CD3<sub>CT</sub>, CD8<sub>CT</sub>, CD57<sub>CT</sub>, CD68<sub>CT</sub>, CD45RO<sub>IM</sub>, CD68<sub>IM</sub>). Thus, the chemokines CXCL9, CXCL10, and CX3CL1 may thus attract different subsets of immune cells to different locations in the tumor. We performed Cox multivariate analysis combining soluble factors and immune cells. Each chemokine was dependent on the density of T cell subsets, whereas the density of T cell subsets remained significant, indicating the relationships between these factors (data not shown). In summary, these data showed that the phenotypes of the intratumoral immune cells are strongly associated with specific chemokines and adhesion molecules, indicating a high degree of functional coordination.

### **Intratumoral T cell repertoire, chemokines and prognosis**

Although total T cell subpopulations correlated with chemokine levels, one might expect that certain T cells, that are selectively attracted to the tumor site, are specific in nature. To test this hypothesis, we analyzed the T cell repertoire of tumor infiltrating T cells for 10 randomly selected patients. All TCR V $\beta$  chain families (V $\beta$ 1-V $\beta$ 24) and 13 different CDR3 lengths for each V $\beta$  were quantified. 3D-density plots representing the

total quantity of each TCR are represented for each patient. The results showed firstly that all V $\beta$  and most CDR3 lengths are present within the tumor of the patients (Figure 4, A and C). They also showed that the amount of each TCR was different for each patient. Notably, the TCR repertoire of patients with a high CX3CL1 level was clearly distinguishable from the repertoire of patients with a low CX3CL1 expression level. An example is illustrated for V $\beta$ 2 (Figure 4, B and D). The levels of V $\beta$ 2 were higher in patients with high CX3CL1 levels than in patients with low CX3CL1 levels. By extending the analysis of the specificity of the V $\beta$ 2 T cells to the CDR3 lengths, it was found that all patients with high CX3CL1 levels possessed significantly increased levels of the V $\beta$ 2L03, V $\beta$ 2L07, V $\beta$ 2L06 and V $\beta$ 2L09 T cell receptors than patients with low CX3CL1 levels ( $P < 0.01$ ,  $P < 0.008$ ,  $P < 0.05$  and  $P < 0.05$ , respectively) (Figure 5B, and data not shown). The other CDR3 lengths for V $\beta$ 2 did not differ significantly between patients. In order to analyze the T cell repertoire in a global manner, we calculated a correlation matrix between all TCR rearrangements together with chemokine expression levels (Figure 5E, Supplementary Table 5). This visualization method revealed six major clusters of specific co-modulated TCR rearrangements. Subsets of specific T cells are thus simultaneously overrepresented in the same patients. The majority of T cells were found in the same cluster (cluster C). Interestingly, CX3CL1, CXCL9 and CXCL10 gene expression levels correlated with specific T cell rearrangements within another cluster (cluster E). This indicated that these chemokines may attract subsets of T cells with a distinct TCR.

We lastly investigated whether the densities of these specific T cells correlated with patient survival. Specific T cells (such as V $\beta$ 2L04, V $\beta$ 5.2L05 or V $\beta$ 5.2L04) found in



cluster C whose intratumoral densities did not differ according to the level of CX3CL1, CXCL9 and CXCL10 gene expression (Figure 5A), were not associated with patient survival (Figure 5C). In contrast, specific T cells found in cluster E or whose density was higher in tumors with high expression levels of CX3CL1, CXCL9, or CXCL10 (such as V $\beta$ 5.2L08, V $\beta$ 2L03 or V $\beta$ 2L07) (Figure 5B) did correlate with patient survival (Figure 5D). Overall survival rates at 3 years were respectively 100% and 28% in patients with tumors with high and low levels of infiltration by V $\beta$ 5.2L08, V $\beta$ 2L03 or V $\beta$ 2L07 T cells. Kaplan-Meier plots displaying the duration of overall survival (7 years follow-up) according to intratumoral T cells densities showed the improved prognosis associated with specific T cells (Figure 5D).

## Discussion

The staggering complexity of multifactorial diseases such as cancer poses significant challenges for the development of stratified or personalized therapies. The integrated analysis of diverse datasets might circumvent these challenges and provide an enhanced understanding of complex systems, such as the tumor microenvironment. We applied such an integrated approach and performed global analyses of the phenome (large-scale flow cytometry experiments), transcriptome, tissue microarrays of specific tumor regions and T cell repertoire analysis, in the tumor microenvironment of patients with colorectal cancer. Our data revealed mechanisms resulting in high or low densities of specific immune cells at the tumor site. Chemokines and adhesion molecules appeared to target immune effector T cells with a specific TCR-repertoire within the tumor. Furthermore the presence within a tumor of T cells with a specific TCR-repertoire correlated with patient survival.

It has been proposed that the limitations of individual studies that are due to experimental design can be overcome by analyzing data obtained from two or more different approaches<sup>23</sup>. Our study, in which we have obtained concordant results using different experimental approaches, demonstrates the potential of this general approach. We provide support for hypothesis-driven research in man using prior knowledge and integrative biology. The prediction of genes associated with prognosis was based on biomolecular network reconstruction using different data sets (conserved genomic neighborhood, phylogenetic profiling, co-expression analysis, protein-protein interaction, functional genomic public database, literature co-occurrence). Interpretation of our data was facilitated by a novel visualization method combining hazard ratio values, a structured description of known biologic information at different levels of granularity

(GO), and tools for data integration and visualization (Cytoscape). The functional patterns of biological markers that we uncovered led us to formulate hypotheses associating specific chemokines with immune cells found at the tumor site. Specifically, predictions for the roles of specific chemokines (CX3CL1, CXCL9, CXCL10, CCL2, CCL5, CCL11) and adhesion molecules (MADCAM1, ICAM1, VCAM1) were experimentally corroborated.

Binding of chemokines and adhesion molecules such as CX3CL1 and MADCAM1 to their receptors (CX3CR1 and A $\beta$ 7 integrin, respectively)<sup>24, 25</sup> controls the recruitment and the adhesion of effector CD8<sup>+</sup> T cells to the intestine. CX3CL1 binds to CD8<sup>+</sup> T cells displaying strong cytotoxicity and expressing Perforin<sup>+</sup>, GZMB<sup>+</sup>, CD57<sup>+</sup>, CD11a<sup>+</sup>, CCR7<sup>-</sup>, CD62L<sup>-</sup>. We showed that the combined expression of CX3CL1 and other predicted chemokines or MADCAM1 induced a significant delay in tumor recurrence. Our data suggest that, in CRC, the majority of cells attracted by CX3CL1 were T<sub>H</sub>1 and effector activated cytotoxic T cells whereas CXCL9 and CXCL10 attracted mostly memory CD45RO<sup>+</sup> T cells. We previously showed that patients without tumor recurrence had higher memory T cell densities than those whose tumors had recurred.<sup>18</sup> Patients with increased densities of CD57<sup>+</sup> or T-Bet<sup>+</sup> cells in each tumor region presented with statistically prolonged disease-free survival. Combinations of effector activated cytotoxic T-cells (CD8, GZM, CD57, T-Bet) and memory T-cells (CD45RO) led to increasing hazard ratio for DFS (data not shown). Several reports analyzed chemokines or adhesion molecules in colorectal cancer. Chemokine receptors (CXCR4, CXCR5, CCR7) and adhesion molecules (CD44, E-cadherin) have been previously associated with tumor invasion and bad prognosis<sup>26, 27</sup>. In two studies, chemokines

(CX3CL1, CXCL16) were associated with good prognosis<sup>28, 29</sup>. However, the power of the reconstructed predictive biomolecular network approach we described herein is that it simultaneously integrates multiple molecules, including specific chemokines and adhesion molecules that could act in a coordinated manner. These results suggested possible mechanisms resulting in high or low densities of specific immune cells in CRC. Thus, the combined expression of chemokines and adhesion molecules - and the resultant density of T cell subsets within the primary tumor - may prevent tumor expansion and recurrence.

The constant genomic metamorphosis of tumor cells<sup>30</sup> may eventually give rise to new phenotypes that display increased or reduced immunogenicity. The subset of effector cytotoxic T cells found within the tumor are however likely to recognize multiple antigens expressed by the tumor cells. Thus, tumor infiltration by cytotoxic and memory T lymphocytes could reflect a level of antitumor immunity shaped by multiple tumor parameters, such as altered expression level of HLA molecules, the expression pattern of tumor antigens, the mutational pathways (MSI, CIMP, CIN)<sup>5, 7</sup>. Several studies previously reported oligoclonal or polyclonal intratumoral T cell repertoire in solid tumors<sup>31</sup>. However, none of these studies analyzed the intratumoral TCR repertoire in relation with chemokine expression or prognosis. First, by analyzing the TCR usage, we clearly found a highly polyclonal intratumoral T cell repertoire since all V $\beta$  and most of the CDR3 lengths were present within the tumor. Second, it can be hypothesized that a majority of T cells are inflammation-related (cluster C) whereas CXCL9, CXCL10, CX3CL1 chemokines attract a subset of specific anti-tumor T cells with a particular

repertoire (cluster E). Third, the presence of a specific intratumoral TCR-repertoire correlated with the survival of the patient.

In summary, we first revealed mechanisms determining the densities (high or low) of specific immune cells in colorectal tumors. Second, chemokines and adhesion molecules appeared to target immune effector T cells with particular TCR-repertoire within the tumor. Furthermore, the presence of a specific intratumoral TCR-repertoire correlated with the survival of the patient. This study also suggests that creation of similar predictive networks could be used to focus biologic research on specific molecules or pathways in broad range of physiologic and pathologic processes.

ACCEPTED MANUSCRIPT

## **Materials and Methods**

### **Patients and database**

The records of colorectal cancer patients who underwent a primary resection of their tumor at the Laennec-HEGP Hospitals between 1996 and 2004 were reviewed and previously described<sup>18</sup>. Histopathological and clinical findings were scored according to the UICC-TNM staging system. For details, see Supplementary Methods and Supplementary Table 6. A secure Web-based database TME.db was built for the management of patients' data. Ethical, Legal and Social Implications were approved by ethical review board.

### **Gene expression analysis**

Frozen tumor samples (cohort 1, n=108; and re-validation cohort 2, n=27) of randomly selected patients available from Laennec-HEGP Hospitals (1996-2004), with sufficient RNA quality and quantity, were selected for gene expression analysis. Total RNA was isolated by homogenization with RNeasy isolation-kit (Qiagen, Valencia, CA). Quantitative real-time TaqMan-PCR was performed using Low-Density-Arrays and the 7900 robotic real-time PCR-system (Applied-Biosystems). Data were analyzed using the SDS Software v2.2 (Applied-Biosystems) and TME.db statistical module.

### **Large-scale flow cytometric analysis**

After mechanical dispersion, cells from fresh tumors were washed and subjected to four-color flow cytometry. Cells were resuspended in PBS/0.5% BSA and incubated for 30 minutes at 4°C with antibodies and relevant isotype controls. 40,000 cells were analyzed per run. Analyses were performed with a FACScalibur flow cytometer and CellQuest software (Becton Dickinson, San Diego, CA).

### **T cell receptor repertoire analysis**

The CDR3 length distribution analysis was done by performing reverse-transcription of V and V–J gene composition and transcripts into cDNA, CDR3-encoding mRNA was amplified by PCR using specific V and C primers. The intratumoral T cell repertoire was performed on 10 randomly selected colorectal tumors using the TcLandscape technology (TcLand, Nantes, France).

### **Tissue Microarray and Immunohistochemistry**

Using a tissue-array instrument (Beecher Instruments, Alphelys, Plaisir, France), two representative regions of the tumor (center (CT) and invasive margin (IM)) were punched from paraffin-embedded tissue-blocks. Tissue-microarray sections were incubated with monoclonal antibodies against CD3 (SP7), CD8 (4B11), CD45RO (OPD4), GZMB (GrB-7), CD57 (NK1), CD1A (O10), cytokeratin (AE1AE3) and cytokeratin-8 (Neomarkers, Fremont, CA), T-bet (4B10) (Santa Cruz Biotechnology, Santa Cruz, CA), and CD68 (PG-M1) (Dako, Copenhagen, Denmark). Envision+ system and DAB-chromogen were applied (Dako). Slides were analyzed using an image analysis workstation (Spot Browser®, Alphelys).

### **Statistical analysis**

The gene prediction network in Fig. 1 was created using the STRING database and GO. Correlation matrix were perform using Pearson un-centered hierarchical clustering. For pairwise comparisons of parametric and non-parametric data the Student's t-test and Wilcoxon rank-sum test were used, respectively. Kaplan Meier estimators of survival were used to visualize the survival curves. Hazard ratio (Cox proportional hazards model) and the logrank test were used to compare disease-free and overall survival between patients in different groups. To avoid over-fitting, hazard ratios obtained by the “minimum p-value” approach were corrected as suggested by Hollander<sup>32</sup>. P-values for

the HiHi vs. LoLo gene combination analysis were corrected for multiple testing using the Benjamini-Hochberg method. We applied the Kruskal-Wallis one-way analysis of variance to determine if any of the patient cohorts was significantly different regarding the clinical parameters; no significant difference was found between cohorts. All through the text a p-value  $< 0.05$  was considered statistically significant. All analyses were performed with the statistical software R (survival package) and Statview. For details, see Supplementary Methods.

ACCEPTED MANUSCRIPT



## Acknowledgements

We are grateful to Dr. Ion Gresser and Dr. John Nelson for providing helpful comments and critical review of the manuscript. This work was supported by grants from the Association pour la Recherche sur le Cancer (ARC), the National Cancer Institute (INCA), the Canceropole Ile de France, Ville de Paris, INSERM, the Austrian Federal Ministry of Science and Research (GEN-AU project Bioinformatics Integration Network), the Austrian Science Fund (SFB project Lipotoxicity), and the European Commission (7FP, Geninca Consortium, grant number 202230).

ACCEPTED MANUSCRIPT

## Figure legends

### Figure 1

Biomolecular network using gene expression data in a cohort of patients with CRC and predicted gene-gene interaction based on available knowledge. The network illustrated experimental data (colored nodes) and *in silico* prediction (white nodes surrounded by a red border). The gene expression data was acquired by a RT-PCR study for 47 genes in a cohort of 108 colorectal cancer (CRC) patients (Supplementary Table 1). The network was reconstructed based on a subset of 12 genes which reached a significant logrank level for disease free survival (DFS). The network shows the top genes predicted *in silico* plus the genes analyzed by RT-PCR. CX3CL1 was the "top one" predicted gene. All nodes surrounded by a red border are predicted by String. The node sizes of the network are based on the hazard ratio for DFS (Supplementary Table 1). Nodes surrounded by a black border had significant logrank  $P$  values ( $P < 0.05$ ). The edge weights of the network are based on the integrated score of the pairwise un-centered Pearson correlation value between the 47 RT-PCR genes and the combined edge scores for all genes predicted *in silico* provided by STRING (see Supplementary methods for details). The network node layout was based on Gene Ontology (GO), gene expression correlations (blue lines), STRING scores (gray lines) and the integrated association strength between genes (edge thickness). Edge thickness levels show the relation strength based on the integrated score value between the nodes. Nodes are colored based on multiple occurrences in different GO categories (Supplementary Table 2).

## Figure 2

Gene expression levels from 108 colorectal tumors (cohort 1) were analyzed by real-time quantitative PCR. (A) Kaplan-Meier curves for the duration of DFS, according to the expression of the predicted genes (CX3CL1, CXCL9, CXCL10, CCL2, CCL5, CCL11, and MADCAM1) were performed. Patients with high (Hi) expression for both genes (red line) or low (Lo) expression for both genes densities (black line), and heterogeneous expression (HiLo or LoHi, green line) are represented. (B) Hazard ratios were calculated for high and low gene expression compared to the whole cohort (108 patients). A hazard ratio-matrix (heatmap) followed by unsupervised hierarchical clustering was represented from favorable prognosis HR=0.4 (red) to poor prognosis HR=2.5 (blue). All HR with HR<0.55 or HR>1.66, were significant.

## Figure 3

(A) Tissue-MicroArray spots of representative tumors with low (top) or high (bottom) CX3CL1, CXCL9 and CXCL10 gene expression are illustrated. Stainings for tumor cells (cytokeratin-8+, blue staining) and cytotoxic T cells (CD8+, brown staining) were performed. (B) Comparison of immune cell densities as measured by flow cytometry from 39 freshly resected tumors. Cytotoxic (CD3+CD8+), memory (CD3+CD45RO+) and TCR $\alpha\beta$  (CD3+TCR $\alpha\beta$ +) T cells were analyzed in the tumors from patients with high - (blue histogram) or low- (white histogram) CX3CL1, CXCL9 or CXCL10 gene expression. (C) Hierarchical clustering of correlation matrix of the flow cytometry data from 39 freshly resected tumors. Pearson correlation coefficients (R) were calculated between the combination of 149 markers for major immune cell populations ("total"

prefix) and specific T cell subpopulations and CX3CL1, CXCL9, and CXCL10. Correlation coefficients were plotted with negative correlation (green), positive correlation (red), and R=0 (yellow), in matrix representation followed by unsupervised Spearman hierarchical clustering. Six major clusters (A-F) are illustrated. (D) Comparison of the mean of immune cell densities (cell/mm<sup>2</sup>) as measured by tissue-micro-arrays from 108 paraffin-embedded tumors. Cytotoxic (CD8+, GZMB+), Th1 (T-Bet+) and memory (CD45RO+) T cells were analyzed in the tumors from patients with high - (blue histogram) or low- (white histogram) CX3CL1, CXCL9 or CXCL10 gene expression. \*  $P < 0.05$ .

#### Figure 4

T cell repertoire analyses were performed on tumors from 10 randomly selected colorectal cancer patients (A to J) using the quantitative expression of the 26 TCR V $\beta$  chain families. Results were expressed in ratio V $\beta$ /HPRT. These quantitative data were represented by the height of the peak in the TcLandscapes (panels A, C) and on the histograms (panels B, D). T cell repertoire analysis was performed by combining qualitative alterations of V $\beta$  usage at the CDR3 length level (13 different CDR3 length) with the magnitude of expression of each V $\beta$  mRNA species. The CDR3 lengths were represented in the TcLandscapes (panels A, C). (E) Pearson correlation coefficients (R) were calculated between the quantity of all V $\beta$  for each of the 13 different CDR3 lengths and CX3CL1, CXCL9, and CXCL10 gene expression. Correlation coefficients were plotted with negative correlation (green), positive correlation (red), and R=0 (yellow), in

matrix representation followed by unsupervised spearman hierarchical clustering. Six major clusters (A-F) were represented on the matrix.

#### Figure 5

The CDR3 lengths of the TCRs were represented in on the histograms (panels *A*, *B*). (*C*, *D*) Kaplan-Meier curves illustrate the overall survival of patients according to particular TCR expression at the median of the dataset. Panel (*C*) illustrates three TCR ( $V\beta 2L04$ ,  $V\beta 5.2L05$  or  $V\beta 5.2L04$ ) not associated with CX3CL1 gene expression, and panel (*D*) illustrates three TCR ( $V\beta 5.2L08$ ,  $V\beta 2L03$  and  $V\beta 2L07$ ) significantly increased in patients with high expression of CX3CL1.

## References

1. Benoist C, Germain RN, Mathis D. A plaidoyer for 'systems immunology'. *Immunol Rev* 2006;210:229-34.
2. Oltvai ZN, Barabasi AL. Systems biology. Life's complexity pyramid. *Science* 2002;298:763-4.
3. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57-70.
4. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61:759-67.
5. Banerjea A, Bustin SA, Dorudi S. The immunogenicity of colorectal cancers with high-degree microsatellite instability. *World J Surg Oncol* 2005;3:26.
6. Finn OJ. Human tumor antigens, immunosurveillance, and cancer vaccines. *Immunol Res* 2006;36:73-82.
7. Finn OJ. Cancer immunology. *N Engl J Med* 2008;358:2704-15.
8. Coussens LM, Werb Z. Inflammation and cancer. *Nature* 2002;420:860-7.
9. Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. *Annu Rev Immunol* 2004;22:329-60.
10. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, Schreiber RD. IFN $\gamma$  and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* 2001;410:1107-11.
11. Smyth MJ, Thia KY, Street SE, Cretney E, Trapani JA, Taniguchi M, Kawano T, Pelikan SB, Crowe NY, Godfrey DI. Differential tumor surveillance by natural killer (NK) and NKT cells. *J Exp Med* 2000;191:661-8.

12. Koebel CM, Vermi W, Swann JB, Zerafa N, Rodig SJ, Old LJ, Smyth MJ, Schreiber RD. Adaptive immunity maintains occult cancer in an equilibrium state. *Nature* 2007;450:903-7.
13. Atreya I, Neurath MF. Immune cells in colorectal cancer: prognostic relevance and therapeutic strategies. *Expert Rev Anticancer Ther* 2008;8:561-72.
14. Clemente CG, Mihm MC, Jr., Bufalino R, Zurrada S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. *Cancer* 1996;77:1303-10.
15. Diederichsen AC, Hjelmberg JB, Christensen PB, Zeuthen J, Fenger C. Prognostic value of the CD4+/CD8+ ratio of tumour infiltrating lymphocytes in colorectal cancer and HLA-DR expression on tumour cells. *Cancer Immunol Immunother* 2003;52:423-8.
16. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, Makrigiannakis A, Gray H, Schlienger K, Liebman MN, Rubin SC, Coukos G. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* 2003;348:203-13.
17. Pages F, Berger A, Camus M, Sanchez-Cabo F, Costes A, Molidor R, Mlecnik B, Kirilovsky A, Nilsson M, Damotte D, Meatchi T, Bruneval P, Cugnenc PH, Trajanoski Z, Fridman WH, Galon J. Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* 2005;353:2654-66.
18. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, Tosolini M, Camus M, Berger A, Wind P, Zinzindohoue F, Bruneval P, Cugnenc PH, Trajanoski Z, Fridman WH, Pages F. Type, density, and location of immune

- cells within human colorectal tumors predict clinical outcome. *Science* 2006;313:1960-4.
19. Galon J, Fridman WH, Pages F. The adaptive immunologic microenvironment in colorectal cancer: a novel perspective. *Cancer Res* 2007;67:1883-6.
  20. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* 2005;33:D433-7.
  21. Garcia O, Saveanu C, Cline M, Fromont-Racine M, Jacquier A, Schwikowski B, Aittokallio T. GOlorize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. *Bioinformatics* 2007;23:394-6.
  22. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 2009;25:1091-3.
  23. Vidal M. A biological atlas of functional maps. *Cell* 2001;104:333-9.
  24. Berlin C, Berg EL, Briskin MJ, Andrew DP, Kilshaw PJ, Holzmann B, Weissman IL, Hamann A, Butcher EC. Alpha 4 beta 7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1. *Cell* 1993;74:185-95.
  25. Imai T, Hieshima K, Haskell C, Baba M, Nagira M, Nishimura M, Kakizaki M, Takagi S, Nomiyama H, Schall TJ, Yoshie O. Identification and molecular characterization of fractalkine receptor CX3CR1, which mediates both leukocyte migration and adhesion. *Cell* 1997;91:521-30.



26. Gunther K, Leier J, Henning G, Dimmler A, Weissbach R, Hohenberger W, Forster R. Prediction of lymph node metastasis in colorectal carcinoma by expression of chemokine receptor CCR7. *Int J Cancer* 2005;116:726-33.
27. Schimanski CC, Schwald S, Simiantonaki N, Jayasinghe C, Gonner U, Wilsberg V, Junginger T, Berger MR, Galle PR, Moehler M. Effect of chemokine receptors CXCR4 and CCR7 on the metastatic behavior of human colorectal cancer. *Clin Cancer Res* 2005;11:1743-50.
28. Hojo S, Koizumi K, Tsuneyama K, Arita Y, Cui Z, Shinohara K, Minami T, Hashimoto I, Nakayama T, Sakurai H, Takano Y, Yoshie O, Tsukada K, Saiki I. High-level expression of chemokine CXCL16 by tumor cells correlates with a good prognosis and increased tumor-infiltrating lymphocytes in colorectal cancer. *Cancer Res* 2007;67:4725-31.
29. Ohta M, Tanaka F, Yamaguchi H, Sadanaga N, Inoue H, Mori M. The high expression of Fractalkine results in a better prognosis for colorectal cancer patients. *Int J Oncol* 2005;26:41-7.
30. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, Szabo S, Buckhaults P, Farrell C, Meeh P, Markowitz SD, Willis J, Dawson D, Willson JK, Gazdar AF, Hartigan J, Wu L, Liu C, Parmigiani G, Park BH, Bachman KE, Papadopoulos N, Vogelstein B, Kinzler KW, Velculescu VE. The consensus coding sequences of human breast and colorectal cancers. *Science* 2006;314:268-74.
31. Matsutani T, Shiiba K, Yoshioka T, Tsuruta Y, Suzuki R, Ochi T, Itoh T, Musha H, Mizoi T, Sasaki I. Evidence for existence of oligoclonal tumor-infiltrating

lymphocytes and predominant production of T helper 1/T cytotoxic 1 type cytokines in gastric and colorectal tumors. *Int J Oncol* 2004;25:133-41.

32. Hollander N, Sauerbrei W, Schumacher M. Confidence intervals for the effect of a prognostic factor after selection of an 'optimal' cutpoint. *Stat Med* 2004;23:1701-13.

ACCEPTED MANUSCRIPT

Table 1

Mediane immune cell densities according to chemokine expression levels

Cells	Genes	CCL5-Hi	CCL5-Lo	P-value	CXCL9-Hi	CXCL9-Lo	P-value	CXCL10-Hi	CXCL10-Lo	P-value	CX3CL1-Hi	CX3CL1-Lo	P-value
CD3-CT		416.5	188	0.0007 **	367.8	188	0.0094 **	380.5	191.6	0.0046 **	377	188	0.0163 *
CD3intra-CT		42	13	0.3912	52	10	0.0992 (*)	42	10	0.3015	53.5	10	0.0359 *
T.Bet-CT		2.6	1.6	0.2976	4.6	0	0.0265 *	2	1.6	0.4149	6.5	0	0.0020 **
T.Bet-IM		2.5	0	0.5264	3.9	0	0.4081	3.9	0	0.3282	5.3	0	0.0315 *
GZMB-CT		44.5	10.7	0.0539 (*)	40	7.3	0.1566	40	22	0.4298	52.7	0	0.0011 **
GZMB-IM		124	63.4	0.1014	83	71.3	0.3000	124	71.3	0.1713	139	19.5	0.0036 **
CD57-CT		44.1	14.8	0.0025 **	53.1	15.8	0.0027 **	49.3	13.7	0.0013 **	40.4	15.9	0.0020 **
CD8-CT		166.2	46	0.0004 **	134	46.7	0.0042 **	134	37.6	0.0013 **	118.4	53.8	0.0542 (*)
CD45RO-CT		416	135.5	<0.0001 **	366.7	136.5	0.0025 **	380.2	141.8	0.0059 **	282	143.9	0.3483
CD68-CT		921.3	323.3	<0.0001 **	914.6	375.7	0.0001 **	864.2	329.3	<0.0001 **	613.2	438.1	0.0967 (*)
CD1a-CT		3.1	2	0.7600	3.2	1.8	0.2560	3.1	1.8	0.8923	3.4	1.8	0.3298

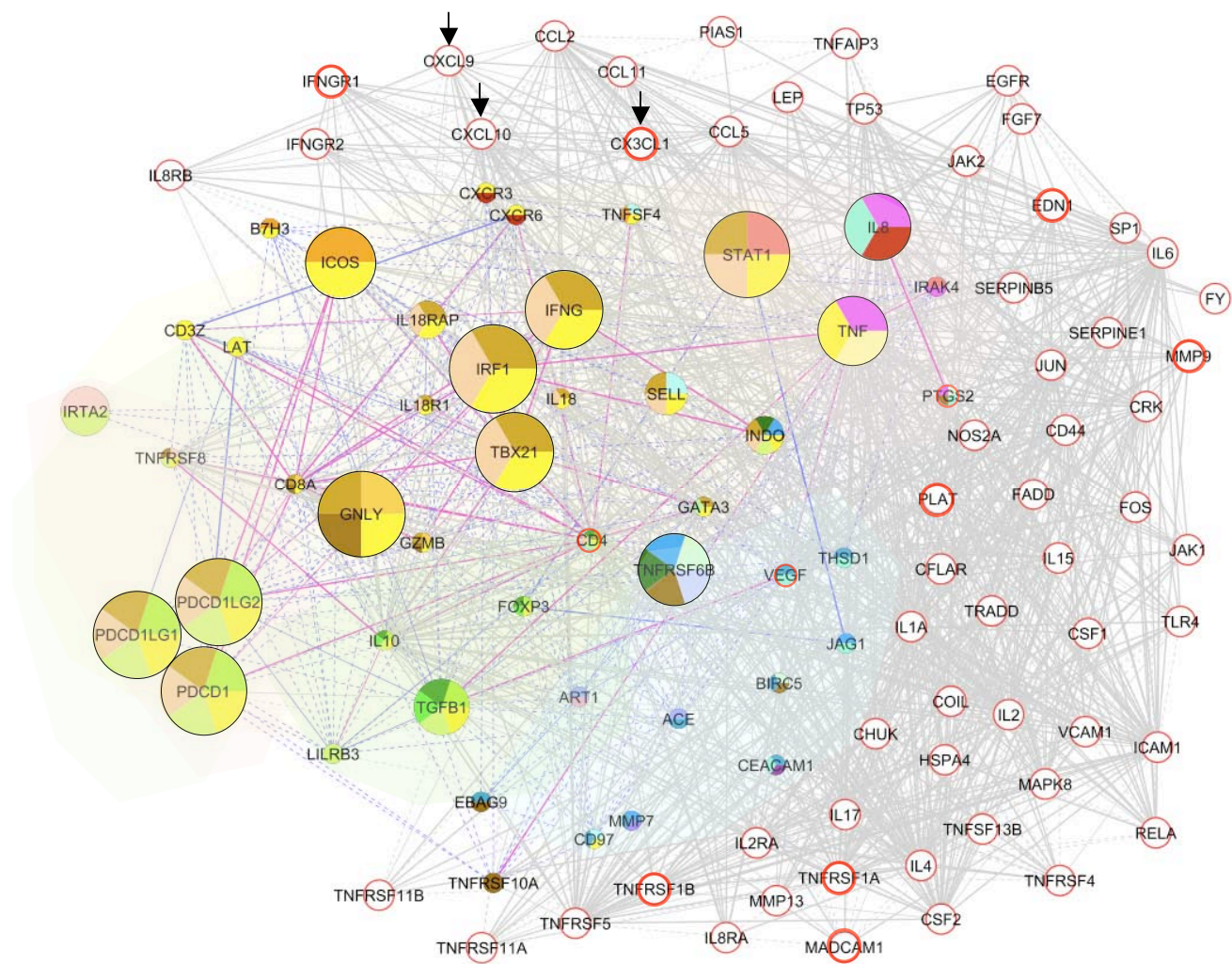


Figure 1

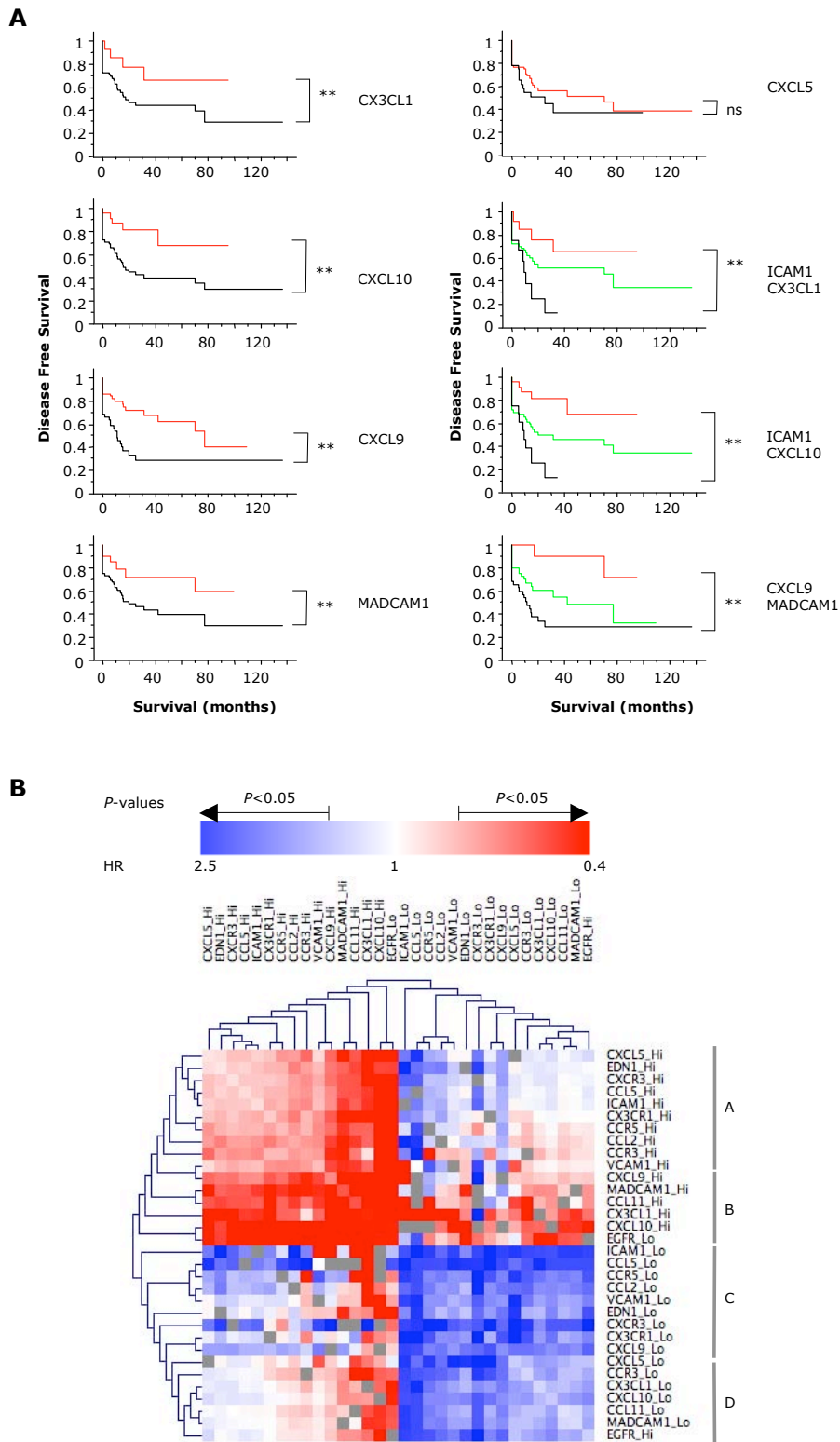


Figure 2

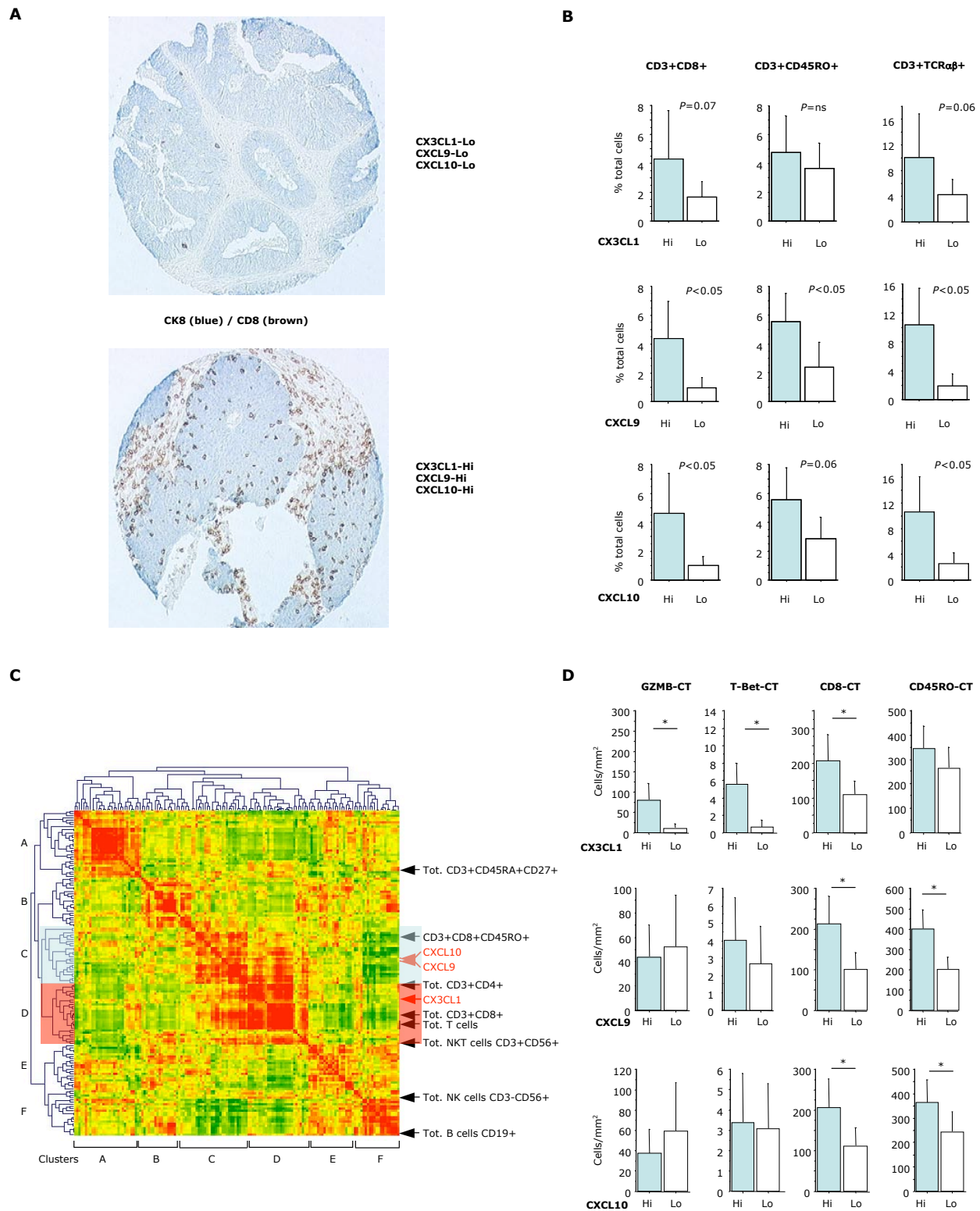


Figure 3

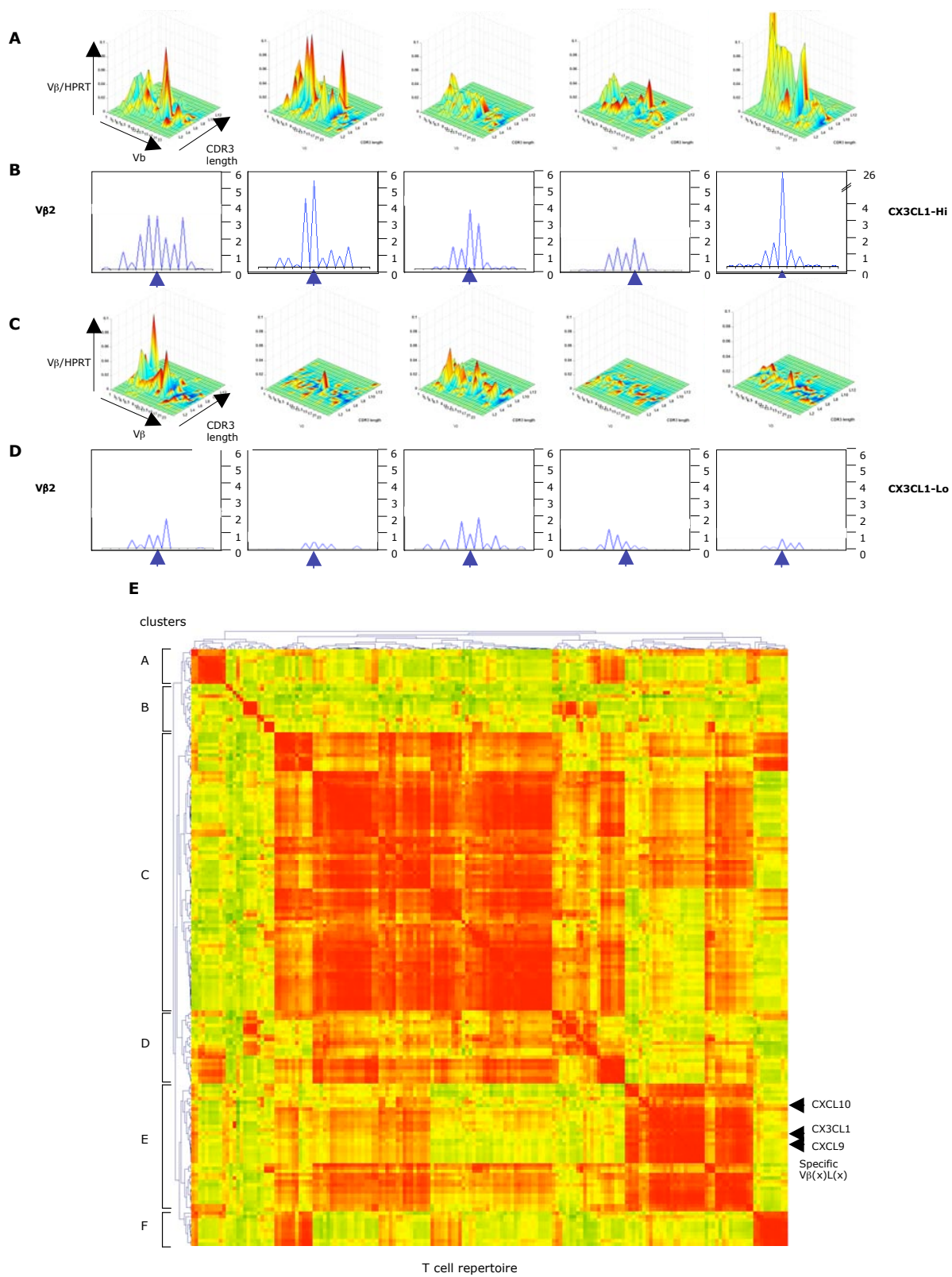


Figure 4

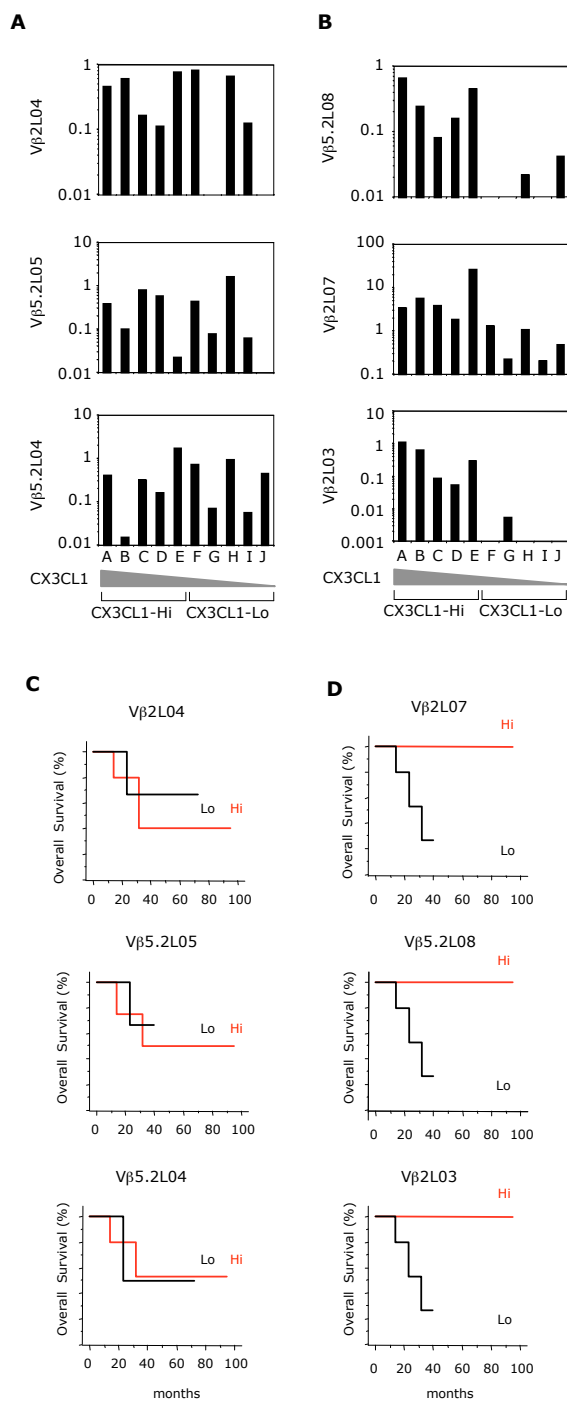


Figure 5



## Supplementary Online Information for

**Biomolecular network reconstruction identifies T cell homing factors associated with survival in colorectal cancer**

Bernhard Mlecnik<sup>1,2,3,4</sup>, Marie Tosolini<sup>1,2,3</sup>, Pornpimol Charoentong<sup>4</sup>, Amos Kirilovsky<sup>1,2,3</sup>, Gabriela Bindea<sup>1,2,3,4</sup>, Anne Berger<sup>5</sup>, Matthieu Camus<sup>1,2,3</sup>, Mélanie Gillard<sup>1,2,3</sup>, Patrick Bruneval<sup>6</sup>, Wolf-Herman Fridman<sup>1,2,3,7</sup>, Franck Pagès<sup>1,2,3,7</sup>, Zlatko Trajanoski<sup>4</sup>, Jérôme Galon<sup>1,2,3,5, #</sup>

This file contains:

Supplementary figure legends 1 to 3

Supplementary Figures 1 to 3

Supplementary Tables 1 to 6

Supplementary experimental procedures

**Supplementary Figure 1**

The chart shows the percentage of the genes found for different GO categories reported to the “top 65” predicted genes (Fig.1). We mapped the genes in Gene Ontology under several selection criteria. GO terms from biological process, between levels 4 to 14, were used. The related GO terms (parent-child relation) with similar associated genes were fused in order to avoid the redundancy. Finally, based on the similarity of their parents, the selected GO terms were associated in functional groups and displayed with different colors in a bar chart. Among the predicted genes were molecules involved in leukocytes and myeloid cell differentiation, in the regulation of apoptosis, in the kinase cascade and in chemotaxis.

**Supplementary Figure 2**

Gene expression levels from a validation series of 27 independent randomly selected colorectal tumors (cohort 2) were analyzed by real-time quantitative PCR. **(A)** Kaplan-Meier curves for the duration of DFS, according to the expression of the predicted chemokines (CX3CL1, CXCL9, and CXCL10). **(B)** Kaplan-Meier curves for the duration of DFS, according to the combined expression of genes (CX3CL1/ICAM1, CXCL10/ICAM1, and CXCL9/MADCAM1). Patients with high (Hi) expression for both genes (red line) or low (Lo) expression for both genes densities (black line), and heterogeneous expression (Het: HiLo or LoHi, blue line) are represented.

**Supplementary Figure 3**

Hierarchical clustered wilcoxon *P*-value matrix (genes vs. TMA) corresponding to the network. Dark yellow and light yellow squares indicate a significant wilcoxon *P*-value < 0.05 and < 0.005, respectively. Details are presented in Table S4.

NAME	UniqueID	Hazard ratio	(95% inf-sup)	Log Rank P-value	Significant
PDCD1LG1	NM_014143	2.82	(1.557-5.092)	0.0003	s
IRF1	NM_002198	2.79	(1.536-5.05)	0.0003	s
GNLY	NM_006433	2.71	(1.511-4.878)	0.0004	s
PDCD1LG2	NM_025239	2.36	(1.326-4.203)	0.0029	s
PDCD1	NM_005018	2.11	(1.188-3.755)	0.0107	s
STAT1	NM_007315	2.03	(1.144-3.608)	0.0121	s
IFNG	NM_000619	1.97	(1.08-3.594)	0.0240	s
TBX21	NM_013351	1.95	(1.103-3.441)	0.0232	s
ICOS	NM_012092	1.9	(1.079-3.339)	0.0349	s
TNFRSF6B	NM_032945	1.87	(1.066-3.292)	0.0332	s
TNF	NM_000594	1.84	(1.041-3.25)	0.0358	s
IL8	NM_000584	1.77	(1.006-3.115)	0.0410	s
TGFB1	NM_000660	1.73	(0.986-3.051)	0.0548	
IRTA2	NM_031281	1.69	(0.964-2.954)	0.0631	
IL18RAP	NM_003853	1.66	(0.949-2.908)	0.0777	
INDO	NM_002164	1.64	(0.94-2.871)	0.0769	
SELL	NM_000655	1.63	(0.927-2.866)	0.0713	
IL18R1	NM_003855	1.6	(0.915-2.785)	0.1052	
MMP7	NM_002423	1.5	(0.859-2.626)	0.1414	
LILRB3	NM_006864	1.49	(0.855-2.602)	0.1448	
PTGS2	NM_000963	1.47	(0.843-2.547)	0.1928	
TNFRSF8	NM_001243	1.44	(0.829-2.515)	0.2148	
CD4	NM_000616	1.43	(0.821-2.496)	0.2079	
CXCR3	NM_001504	1.4	(0.807-2.438)	0.2390	
GAPD	NM_002046	1.39	(0.792-2.423)	0.2428	
CD8A	NM_001768	1.38	(0.79-2.412)	0.2469	
IRAK4	NM_016123	1.38	(0.792-2.392)	0.2687	
CD3Z	NM_198053	1.34	(0.769-2.321)	0.3419	
EBAG9	NM_004215	1.34	(0.765-2.338)	0.2986	
IL10	NM_000572	1.33	(0.766-2.313)	0.3226	
GATA3	NM_002051	1.32	(0.756-2.293)	0.3401	
LAT	NM_014387	1.32	(0.759-2.297)	0.3025	
IL18	NM_001562	1.3	(0.743-2.269)	0.3118	
VEGF	NM_003376	1.27	(0.732-2.212)	0.3505	
THSD1	NM_018676	1.23	(0.708-2.147)	0.4940	
FOXP3	NM_014009	1.23	(0.705-2.132)	0.4515	
CXCR6	NM_006564	1.22	(0.704-2.126)	0.5240	
JAG1	NM_000214	1.22	(0.701-2.12)	0.4693	
TNFRSF10A	NM_003844	1.21	(0.7-2.106)	0.5215	
ACE	NM_000789	1.17	(0.674-2.039)	0.5064	
B7H3	NM_025240	1.17	(0.675-2.037)	0.5276	
BIRC5	NM_001168	1.17	(0.672-2.037)	0.5896	
CD97	NM_078481	1.14	(0.657-1.973)	0.7064	
GZMB	NM_004131	1.13	(0.597-2.137)	0.7527	
CEACAM1	NM_001712	1.05	(0.602-1.815)	0.8513	
TNFSF4	NM_003326	1.03	(0.595-1.79)	0.9210	
ART1	NM_004314	1.03	(0.361-2.942)	0.9149	

Supplementary Table 1

**Biological function of the nodes based on Gene Ontology (GO)**

Function	Class description	GO ID
T cell activation	Adaptive immune response	2250
	T cell activation	42110
	Regulation of T cell mediated cytotoxicity	1914
	Memory T cell differentiation	43379
	T cell costimulation	31295
T cell differentiation	T-helper 1 cell differentiation	45063
	T-helper 2 cell differentiation	45064
Negative regulation of immune response	Negative regulation of lymphocyte activation	51250
	Regulation of regulatory T cell differentiation	45589
	Immune response-inhibiting signal transduction	2765
	Negative regulation of regulatory T cell differentiation	45590
B cell activation	B cell activation	42113
Innate immune response	Innate immune response	45087
Inflammatory response	Inflammatory response	6954
Migration	Chemotaxis	6935
	Leukocyte migration during immune response	2522
	Cell migration	16477
Tumor related	Apoptosis	6915
	Anti-apoptosis	6916
	FasL biosynthesis	45210
	Angiogenesis	1525
	Selectin	8337
	ADP-ribosylation factor binding	30306
	Regulation of coenzyme metabolism	51196
	Metalloendopeptidase activity	4222
	Cell adhesion	7155
	Immune response to tumor cell	2418

**Supplementary Table 2**

NAME	STRING with all parameters		STRING without Experiments and Coexpression	
	Score	Prediction Ranking	Score	Prediction Ranking
CX3CL1	0,999	1	0,999	1
CD4	0,999	2	0,999	2
IFNGR1	0,999	3	0,994	65
TNFRSF1B	0,999	4	0,999	3
EDN1	0,999	5	0,999	4
TNFRSF1A	0,999	6	0,999	5
PTGS2	0,999	7	0,999	6
JAK2	0,999	8	0,999	7
MADCAM1	0,999	9	0,999	8
MMP9	0,999	10	0,999	9
PLAT	0,999	11	0,999	10
SERPINE1	0,999	12	0,999	11
CCL2	0,999	13	0,999	12
IL2	0,999	14	0,999	13
IL17	0,999	15	0,999	14
TNFAIP3	0,999	16	0,999	15
PIAS1	0,999	17	0,996	59
IL6	0,999	18	0,999	16
IL1A	0,999	19	0,999	17
ICAM1	0,999	20	0,999	18
TP53	0,999	21	0,999	19
TNFRSF11A	0,999	22	0,999	20
EGFR	0,999	23	0,999	21
CD44	0,999	24	0,999	22
TNFRSF5	0,999	25	0,999	23
IL2RA	0,999	26	0,999	24
IFNGR2	0,999	27	0,998	51
CCL5	0,999	28	0,999	25
VCAM1	0,999	29	0,999	26
FY	0,999	30	X	X
IL8RA	0,999	31	0,999	27
CSF2	0,999	32	0,999	28
TNFRSF11B	0,999	33	0,999	29
NOS	0,999	34	0,999	30
CRK	0,999	35	0,999	31
FADD	0,999	36	0,999	32
CCL11	0,999	37	0,999	33
CXCL10	0,999	38	0,999	34
RELA	0,999	39	0,999	35
LEP	0,999	40	0,999	36
JUN	0,999	41	0,999	37
VEGF	0,999	42	0,999	38
IL8RB	0,999	43	0,999	39
NOS2A	0,999	44	0,999	40
CSF1	0,999	45	0,999	41
SP1	0,999	46	0,999	42
IL15	0,999	47	0,999	43
MMP13	0,999	48	0,999	44
TRADD	0,999	49	0,999	45
JAK1	0,999	50	0,998	55
CFLAR	0,999	51	0,999	46
TNFSF13B	0,998	52	0,997	57
IL4	0,998	53	0,998	47
COIL	0,998	54	0,998	48
IL10	0,998	55	0,998	49
IL18	0,998	56	0,998	50
HSPA4	0,998	57	0,998	52
TNFRSF4	0,998	58	0,998	53
TLR4	0,998	59	0,998	54
MAPK8	0,998	60	0,998	56
CHUK	0,997	61	0,997	58
CXCL9	0,996	62	0,996	60
SERPIN5	0,996	63	0,996	61
FOS	0,996	64	0,996	62
FGF7	0,995	65	0,995	63
TLR5	X	X	0,995	64

Supplementary Table 3

Table S3: Immune cell densities according to chemokine expression levels

Gene	Tissue MicroArray Marker tested Marker	Center of Tumor region (CT)					Invasive Margin region (IM)				
		High (HI)		Low (Lo)		p-value	High (HI)		Low (Lo)		p-value
		nb.	median (inf.-sup.)	nb.	median (inf.-sup.)		nb.	median (inf.-sup.)	nb.	median (inf.-sup.)	
CCL11	CD3	39	325.7 (41-1372)	30	236 (13-858.1)	0.1263	38	601.1 (62.3-1280)	29	559.9 (111.6-1474.6)	0.6373
CCL11	CD3intra	19	52 (0-156)	10	8.5 (0-94)	0.1330	18	16 (0-77)	9	24 (0-181)	0.5519
CCL11	T.Bet	19	1.8 (0-9.8)	10	2.6 (0-9.7)	0.7934	13	1 (0-13.8)	8	1 (0-14.6)	0.9694
CCL11	GZMB	19	22 (0-99.8)	9	24.9 (0-167.5)	0.8042	14	82.7 (7.5-403)	8	102.2 (3.9-172)	0.9203
CCL11	CD57	41	27 (0-116.5)	36	46.3 (1.7-180.4)	0.1124	37	60.7 (0-252.2)	34	79.5 (0.9-292.3)	0.4672
CCL11	CD8	40	83 (5.9-473)	38	84 (5-377)	0.8376	34	146.5 (8.6-471)	31	143.1 (3-707.6)	0.8194
CCL11	CD8intra	19	81 (0-178)	11	28.8 (0-232)	0.3217	12	86.8 (0-394)	8	73.6 (0-1250)	0.9383
CCL11	CD45RO	40	226.3 (26.6-1004.1)	34	258.5 (4-793)	0.9194	40	244.4 (32-827.3)	34	339.4 (27-1152.7)	0.2365
CCL11	CD68	41	438.1 (45.8-1256.5)	37	927.9 (52.8-3270.5)	0.0014 **	40	1222.8 (58.1-3012.5)	36	1284.4 (107.8-4435.5)	0.5799
CCL11	CD68intra	41	39.6 (3.6-154)	37	77.1 (0.5-232.2)	0.0507 (*)	40	60.4 (0-275.5)	36	47.1 (0-224.7)	0.4110
CCL11	CD1a	41	1.8 (0-18.4)	38	2.6 (0-19.6)	0.5119	38	4.1 (0-24.3)	33	1.8 (0-10.8)	0.1113
CCL2	CD3	33	326.6 (69.2-639.7)	36	215.9 (13-1372)	0.3358	32	566.6 (62.3-1474.6)	35	559.9 (111.6-1310)	0.7675
CCL2	CD3intra	14	39 (0-156)	15	10 (0-126)	0.4160	13	13 (0-52)	14	44 (0-114)	0.1184
CCL2	T.Bet	15	1.7 (0-9.7)	14	1.8 (0-16.8)	0.4149	9	1 (0-5.3)	12	1 (0-14.6)	0.9700
CCL2	GZMB	14	32.5 (0-167.5)	14	12.5 (0-71.3)	0.2973	10	103.2 (7.5-280)	12	77.2 (3.9-219)	0.5824
CCL2	CD57	37	34.4 (1.6-85.2)	40	33.3 (0-180.4)	0.8265	32	61 (1.9-251.5)	39	62.4 (0-351.8)	0.5618
CCL2	CD8	39	104.7 (5.9-377)	39	57.9 (5-473)	0.2917	32	185.2 (3-747)	33	143 (7.5-520.5)	0.2474
CCL2	CD8intra	15	71.8 (0-232)	15	31.3 (0-178)	0.5471	10	86.8 (0-242)	10	70.6 (0-394)	0.8094
CCL2	CD45RO	37	219 (23.8-780)	37	268.5 (4-900)	0.5835	36	330.1 (32-1152.7)	38	297 (27-837.3)	0.4164
CCL2	CD68	38	582.5 (78.6-1948.6)	40	563.1 (45.8-2190.6)	0.6296	38	1171.3 (58.1-3546.2)	38	1284.4 (107.8-2953.8)	0.8162
CCL2	CD68intra	38	63.7 (5-232.2)	40	41.1 (0.5-164.7)	0.2888	38	50.3 (0-275.5)	38	62.9 (0-241.3)	0.3970
CCL2	CD1a	39	3.2 (0-18.4)	40	1.7 (0-19.6)	0.4590	36	2 (0-15.9)	35	4.4 (0-22.1)	0.1694
CCL5	CD3	36	416.5 (14-1372)	33	188 (13-415.9)	0.0007 **	35	562.6 (62.3-1474.6)	32	573 (111.6-1310)	0.5960
CCL5	CD3intra	15	42 (0-156)	14	13 (0-95)	0.3912	14	16 (0-172)	13	21 (0-114)	0.8072
CCL5	T.Bet	14	2.6 (0-9.8)	15	1.6 (0-3.6)	0.2976	10	2.5 (0-40.5)	11	0 (0-14.6)	0.5264
CCL5	GZMB	15	44.5 (0-220.9)	13	10.7 (0-49.9)	0.0539 (*)	11	124 (38.6-280)	11	63.4 (3.9-172)	0.1014
CCL5	CD57	39	44.1 (1.7-158.8)	38	14.8 (0-94.2)	0.0025 **	35	97.8 (9.9-281.8)	36	45.3 (0-239.5)	0.0395 *
CCL5	CD8	42	166.2 (17.3-675)	36	46 (5-286)	0.0004 **	35	236.7 (8.6-747)	30	143.1 (3-471)	0.0662 (*)
CCL5	CD8intra	15	130 (0-479)	15	49 (0-138)	0.1839	11	98.9 (0-242)	9	35.1 (0-143)	0.4457
CCL5	CD45RO	39	416 (4-1137.1)	35	135.5 (14.5-524)	0.0000 **	38	404.7 (32-1152.7)	36	244.4 (27-773.2)	0.0399 *
CCL5	CD68	40	921.3 (192.9-3270.5)	38	323.3 (45.8-736.1)	0.0000 **	40	1471.4 (58.1-3546.2)	36	1046.3 (107.8-2953.3)	0.1790
CCL5	CD68intra	40	79.5 (6.3-232.2)	38	35.7 (0.5-108)	0.0015 **	40	54.6 (0-309.5)	36	54.6 (0-244.8)	0.4382
CCL5	CD1a	41	3.1 (0-28)	38	2 (0-19.6)	0.7600	37	2.9 (0-15.9)	34	4.1 (0-24.3)	0.5083
CCR3	CD3	40	339.8 (14-1372)	29	191.6 (13-639.7)	0.0136 *	38	510.9 (143-1287.4)	29	624.8 (62.3-1440)	0.4540
CCR3	CD3intra	23	36 (0-156)	6	5 (0-94)	0.2435	21	15 (0-77)	6	42.5 (0-114)	0.6600
CCR3	T.Bet	22	1.8 (0-16.8)	7	1.8 (0-9.7)	0.6154	15	1.9 (0-13.8)	6	0 (0-14.6)	0.6504
CCR3	GZMB	22	23.5 (0-99.8)	6	30.3 (0-111.1)	0.8214	17	83 (9-280)	5	63.4 (12.4-172)	0.3396
CCR3	CD57	38	34.8 (0-140.5)	39	31.3 (0-158.8)	0.8147	37	65.1 (0-268)	34	57.4 (4-251.5)	0.8146
CCR3	CD8	40	82.5 (5-473)	38	94.6 (6.3-377)	0.8729	30	210 (7.5-632.2)	35	143 (3-560.9)	0.1132
CCR3	CD8intra	23	71.8 (0-178)	7	7.8 (3.1-232)	0.4615	14	86.8 (0-394)	6	89.1 (0-1250)	0.9670
CCR3	CD45RO	38	281.7 (4-900)	36	167.2 (14.5-705.6)	0.3244	38	224 (27-661.4)	36	339.4 (30-945)	0.2130
CCR3	CD68	39	573.7 (52.8-2190.6)	39	560.6 (45.8-2165.2)	0.7057	39	1340.5 (58.1-3045.6)	37	1115.4 (107.8-3271)	0.3680
CCR3	CD68intra	39	64.6 (4.2-225.4)	39	48.8 (0.5-170.3)	0.9051	39	55.1 (0-275.5)	37	53.8 (0-249.2)	0.9710
CCR3	CD1a	39	3.3 (0-11.3)	40	1.8 (0-22.5)	0.7526	38	4.1 (0-15.9)	33	1.8 (0-22.1)	0.3615
CCR5	CD3	36	426 (69.2-1372)	33	176 (13-415.9)	0.0000 **	36	601.1 (62.3-1474.6)	31	407 (95.8-1280)	0.1369
CCR5	CD3intra	15	25 (0-156)	14	23 (0-95)	0.4955	15	15 (0-114)	12	22.5 (0-56)	0.6236
CCR5	T.Bet	14	4.6 (0-9.8)	15	0 (0-3.6)	0.0602 (*)	9	0 (0-3.9)	12	3.6 (0-40.5)	0.4985
CCR5	GZMB	15	40 (0-220.9)	13	18.8 (0-71.3)	0.2363	11	124 (10.9-280)	11	71.3 (3.9-219)	0.2169
CCR5	CD57	39	48.3 (2.9-158.8)	38	14.3 (0-91.6)	0.0011 **	35	67.7 (0.9-292.3)	36	50.7 (0-239.5)	0.2355
CCR5	CD8	42	182.6 (5-675)	36	32.6 (5.9-226.1)	0.0000 **	35	221 (18-747)	30	141.7 (3-560.9)	0.0518 (*)
CCR5	CD8intra	16	105.5 (0-479)	14	35.6 (0-138)	0.1762	11	98.9 (0-242)	9	81.9 (0-394)	0.8190
CCR5	CD45RO	39	403.7 (23.8-1137.1)	35	135.5 (4-524)	0.0000 **	39	415.8 (39-1152.7)	35	199.2 (27-670.5)	0.0049 **
CCR5	CD68	40	921.3 (110.1-3270.5)	38	324.8 (45.8-1201.2)	0.0000 **	40	1405.2 (58.1-3546.2)	36	1126.3 (107.8-3012.5)	0.3929
CCR5	CD68intra	40	76.7 (6.3-232.2)	38	34.5 (0.5-148.9)	0.0017 **	40	49 (0-275.5)	36	58.5 (0-224.7)	0.4824
CCR5	CD1a	41	3.1 (0-19.5)	38	1.8 (0-19.6)	0.8767	37	2.6 (0-15.9)	34	4.3 (0-27.2)	0.3017
CX3CL1	CD3	37	377 (59.6-970.4)	35	188 (13-358.6)	0.0163 *	35	559.9 (62.3-1440)	35	616.2 (95.8-1287.4)	0.7350
CX3CL1	CD3intra	16	53.5 (0-126)	13	10 (0-68)	0.0359 *	14	35 (0-202)	13	17 (0-88)	0.4494
CX3CL1	T.Bet	16	6.5 (0-16.8)	13	0 (0-3.6)	0.0020 **	11	5.3 (0-13.1)	11	0 (0-1)	0.0315 *
CX3CL1	GZMB	16	52.7 (7.1-220.9)	11	0 (0-24.9)	0.0011 **	12	139 (38.6-403)	11	19.5 (3.9-172)	0.0036 **
CX3CL1	CD57	42	40.4 (0-158.8)	39	15.9 (0-91.6)	0.0020 **	36	100.5 (5.2-315.9)	37	39.7 (0-171.5)	0.0044 **
CX3CL1	CD8	44	118.4 (6.3-473)	38	53.8 (5-337.6)	0.0542 (*)	36	143.1 (7.5-747)	34	164.7 (3-520.5)	0.8199
CX3CL1	CD8intra	16	123 (0-232)	14	36.5 (0-81)	0.1283	10	121 (0-242)	11	42.3 (0-204)	0.3412
CX3CL1	CD45RO	41	282 (23.6-1004.1)	34	143.9 (14.5-793)	0.3483	40	331.5 (24.5-945)	36	246.1 (27-1152.7)	0.6753
CX3CL1	CD68	43	613.2 (38.6-2190.6)	39	438.1 (45.8-1681)	0.0967 (*)	41	1115.4 (58.1-3546.2)	39	1254.7 (133.2-3371)	0.8482
CX3CL1	CD68intra	43	65.6 (6.3-261.9)	39	39.6 (0.5-148.9)	0.1260	41	56.6 (0-248.3)	39	53.3 (0-158.4)	0.3120
CX3CL1	CD1a	44	3.4 (0-19.6)	39	1.8 (0-11.3)	0.3298	39	4.4 (0-35)	36	1.4 (0-13.5)	0.0574 (*)
CX3CR1	CD3	35	325.7 (14-1130)	34	203.4 (13-970.4)	0.0786 (*)	33	427.1 (129.3-1310)	34	645.2 (62.3-1474.6)	0.3396
CX3CR1	CD3intra	18	13 (0-156)	11	55 (0-94)	0.8033	16	14 (0-114)	11	33 (0-202)	0.4129
CX3CR1	T.Bet	18	1.7 (0-9.8)	11	1.8 (0-9.7)	0.5599	12	0 (0-5.3)	9	13.8 (0-297.4)	0.0837 (*)
CX3CR1	GZMB	17	18.8 (0-99.8)	11	35.6 (0-111.1)	0.6019	14	79.9 (3.9-280)	8	134 (12.4-403)	0.3650
CX3CR1	CD57	37	31.3 (1.6-116.5)	40	41.6 (0-180.4)	0.7327	35	57.5 (0.9-143.6)	36	94.8 (0-351.8)	0.1064
CX3CR1	CD8	38	84.5 (5-473)	40	70.5 (6.3-372.9)	0.4444	29	221 (22.5-520.5)	36	138.7 (3-632.2)	0.1456
CX3CR1	CD8intra	19	49 (0-178)	11	116 (0-232)	0.7794	11	26.7 (0-242)	9	226 (0-1250)	0.0731 (*)
CX3CR1	CD45RO	36	311.4 (4-900)	38	150.8 (14.5-705.6)	0.1975	35	229 (27-827.3)	39	231 (32-1152.7)	0.2027
CX3CR1	CD68	38	526.6 (52.8-1908.7)	40	638 (45.8-2857.7)	0.3784	36	1171.3 (58.1-3045.6)	40	1334.8 (139.6-4435.5)	0.1389
CX3CR1	CD68intra	38	58.8 (5-225.4)	40	51.9 (0.5-170.3)	0.8541	36	51.2 (0-221.8)	40	63.7 (0-249.2)	0.3849
CX3CR1	CD1a	37	3.4 (0-18.4)	42	1.8 (0-28)	0.5309	37	4.3 (0-15.9)	34	2.6 (0-20.7)	0.6932
CXCL10	CD3	36	380.5 (42.6-1372)	33	191.6 (13-622)	0.0046 **	36	578.3 (62.3-1474.6)	31	545.7 (95.8-1310)	0.3321
CXCL10	CD3intra	15	42 (0-156)	14	10 (0-89)	0.3015	15	15 (0-77)	12	22.5 (0-114)	0.8636
CXCL10	T.Bet	14	2 (0-9.8)	15	1.6 (0-7.9)	0.4149	9	3.9 (0-40.5)	12	0 (0-14.6)	0.3282
CXCL10	GZMB	15	40 (0-111.1)	13	22 (0-99.9)	0.4298	11	124 (10.9-280)	11	71.3 (3.9-219)	0.1713
CXCL10	CD57	40	49.3 (1.7-139.5)	37	13.7 (0-83.2)	0.0013 **	35	100 (0.9-281.8)	36	45.3 (0-193.3)	0.1497
CXCL10	CD8	43	134 (5-675)	35	37.6 (5.9-312)	0.0013 **	35	221 (7.5-747)	30	141.7 (3-520.5)	0.1360
CXCL10	CD8intra	16	76.4 (0-232)	14	34.3 (0-138)	0.3599	11	98.9 (0-242)	9	35.1 (0-143)	0.3805
CXCL10	CD45RO	40	380.2 (23.8-1137.1)	34	141.8 (4-524)	0.0059 **	40	370.5 (39-1152.7)	34	244.4 (27-773.2)	0.0426 *
CXCL10	CD68	41	864.2 (110.1-3270.5)	37	329.3 (45.8-1275.4)	0.0000 **	40	1503.9 (195.8-3546.2)	36	1046.3 (58.1-2953.3)	0.0365 *
CXCL10	CD68intra	41	76.2 (6.3-261.9)	37	31.8 (0.5-148.9)	0.0037 **	40	52.4 (0-227)	36	55.8 (0-196.1)</	

Table S3: Immune cell densities according to chemokine expression levels

Gene	Tissue MicroArray Marker tested Marker	Center of Tumor region (CT)					Invasive Margin region (IM)				
		High (HI)		Low (Lo)		p-value	High (HI)		Low (Lo)		p-value
		nb.	median (inf.-sup.)	nb.	median (inf.-sup.)		nb.	median (inf.-sup.)	nb.	median (inf.-sup.)	
CXCL5	CD3	39	224.7 (13-625.5)	30	365 (59.6-1372)	0.1822	38	516 (62.3-1474.6)	29	570.5 (111.6-1310)	0.7677
CXCL5	CD3intra	18	26 (0-156)	11	25 (0-89)	0.5114	17	21 (0-60)	10	14 (0-81)	0.8597
CXCL5	T.Bet	18	0.8 (0-2.2)	11	6.2 (0-9.8)	0.0931 (*)	15	0 (0-13.8)	6	1.5 (0-14.6)	0.7108
CXCL5	GZMB	18	20.4 (0-167.5)	10	24.9 (0-44.5)	0.7351	13	124 (7.5-219)	9	77.4 (3.9-133)	0.6470
CXCL5	CD57	43	27 (0-140.5)	34	39.8 (2.9-158.8)	0.2164	39	55.8 (0-193.3)	32	91.6 (0.9-415.2)	0.0883 (*)
CXCL5	CD8	44	65.6 (5.9-377)	34	146 (5-473)	0.1802	35	146 (3-632.2)	30	157.3 (9.5-520.5)	0.9323
CXCL5	CD8intra	19	49 (0-232)	11	130 (0-193)	0.6356	13	42.3 (0-394)	7	143 (4.3-599)	0.2655
CXCL5	CD45RO	43	143.1 (4-780)	31	315.5 (23.6-900)	0.1651	43	244 (32-945)	31	339 (27-1026.4)	0.4074
CXCL5	CD68	43	552.4 (45.8-1681.3)	35	702.4 (52.8-2165.2)	0.1175	42	1219.9 (139.6-3012.5)	34	1284.4 (58.1-3045.6)	0.9545
CXCL5	CD68intra	43	48.8 (0.5-232.2)	35	57.8 (4.2-159.6)	0.6602	42	55.8 (0-309.5)	34	52.7 (0-147.7)	0.2042
CXCL5	CD1a	44	3.3 (0-17.8)	35	1.6 (0-19.6)	0.5456	39	2.9 (0-15.9)	32	3.8 (0-22.1)	0.7977
CXCL9	CD3	36	367.8 (42.6-1372)	33	188 (13-622)	0.0094 **	35	624.8 (62.3-1474.6)	32	476.4 (111.6-1064.6)	0.0614 (*)
CXCL9	CD3intra	15	52 (0-156)	14	10 (0-74)	0.0992 (*)	14	16 (0-202)	13	21 (0-56)	0.5914
CXCL9	T.Bet	14	4.6 (0-9.8)	15	0 (0-3.6)	0.0265 *	9	3.9 (0-14.6)	12	0 (0-13.8)	0.4081
CXCL9	GZMB	15	40 (0-111.1)	13	7.3 (0-49.9)	0.1566 **	11	83 (10.9-280)	11	71.3 (3.9-172)	0.3000
CXCL9	CD57	38	53.1 (1.7-158.8)	39	15.8 (0-91.6)	0.0027 **	34	100.5 (0.9-292.3)	37	45.1 (0-193.3)	0.9417 *
CXCL9	CD8	41	134 (5-675)	37	46.7 (5.9-377)	0.0042 **	35	175.3 (7.5-747)	30	144.5 (3-501)	0.3560
CXCL9	CD8intra	15	130 (0-479)	15	39.8 (0-138)	0.1706 **	11	143 (4.3-242)	9	35.1 (0-91.7)	0.1473
CXCL9	CD45RO	39	366.7 (23.8-1137.1)	35	136.5 (4-524)	0.0025 **	38	370.5 (30-1152.7)	36	245.8 (27-773.2)	0.0949 (*)
CXCL9	CD68	39	914.6 (78.4-3270.5)	39	375.7 (45.8-1275.4)	0.0001 **	39	1470 (58.1-3546.2)	37	1202.2 (133.2-1850.1)	0.1915
CXCL9	CD68intra	39	76.2 (7.9-232.2)	39	31.8 (0.5-164.7)	0.0097 **	39	50.9 (0-269.6)	37	60.4 (0-244.8)	0.6288
CXCL9	CD1a	40	3.2 (0-28)	39	1.8 (0-11.2)	0.2560	36	2.8 (0-20.7)	35	3.9 (0-22.1)	0.7323
CXCR3	CD3	36	308.3 (41-1372)	33	237 (13-1036.8)	0.5228	33	562.6 (62.3-1287.4)	34	563.3 (111.6-1474.6)	0.7132
CXCR3	CD3intra	15	25 (0-116)	14	26 (0-156)	0.5826	12	15 (0-77)	15	21 (0-114)	0.8252
CXCR3	T.Bet	16	1.7 (0-2.9)	13	3.6 (0-16.8)	0.2747	9	3.9 (0-13.8)	12	0.5 (0-14.6)	0.8509
CXCR3	GZMB	15	14.3 (0-99.8)	13	35.6 (0-220.9)	0.1091	11	82.3 (10.9-280)	11	133 (3.9-197)	0.6063
CXCR3	CD57	40	33 (0-139.5)	37	35.3 (0-140.5)	0.3430	36	74.1 (0-252.2)	35	52.5 (1.8-317.5)	0.7018
CXCR3	CD8	41	84.6 (5-473)	37	83.6 (6.3-377)	0.9282	34	145.9 (7.5-707.6)	31	147 (3-560.9)	0.8808
CXCR3	CD8intra	15	31.3 (0-178)	15	49.8 (3.1-479)	0.1146	10	58.5 (0-394)	10	117.4 (0-1250)	0.6489
CXCR3	CD45RO	39	268.5 (23.6-1137.1)	35	177.5 (4-793)	0.5393	38	322.9 (30-1026.4)	36	257 (27-945)	0.7107
CXCR3	CD68	41	560.6 (45.8-1634.7)	37	614.2 (52.8-2190.6)	0.7577	40	1252.2 (58.1-4435.5)	36	1224.4 (133.2-2953.3)	0.9629
CXCR3	CD68intra	41	40.8 (3.5-170.3)	37	64.6 (0.5-261.9)	0.1945	40	54 (0-224.7)	36	58.5 (0-275.5)	0.6964
CXCR3	CD1a	42	3.3 (0-19.6)	37	1.8 (0-18.4)	0.5342	36	4.4 (0-27.2)	35	1.8 (0-15.2)	0.1033
EDN1	CD3	38	321.9 (13-1036.8)	31	224.7 (42.6-1161)	0.9952	35	556 (111.6-1310)	32	578.3 (62.3-1287.4)	0.8079
EDN1	CD3intra	19	10 (0-156)	10	42 (0-126)	0.4456	16	15 (0-55)	11	56 (0-202)	0.1647
EDN1	T.Bet	18	1.8 (0-6.8)	11	1.8 (0-3.6)	0.7090	12	2.7 (0-131.5)	9	1 (0-3.9)	0.6249
EDN1	GZMB	19	24.9 (0-167.5)	9	14.3 (0-71.3)	0.6375	12	74.4 (7.5-403)	10	128.5 (3.9-219)	0.7223
EDN1	CD57	41	35.8 (0-140.5)	36	30 (0-180.4)	0.8064	38	74.1 (0.9-292.3)	33	57.5 (0-286)	0.2677
EDN1	CD8	43	84.4 (5-372.9)	35	80.1 (9.1-675)	0.6989	34	201.9 (21.5-747)	31	143 (3-512)	0.1139
EDN1	CD8intra	19	28.8 (0-232)	11	81 (7.6-178)	0.1209	12	34.5 (0-143)	8	215 (0-599)	0.1412
EDN1	CD45RO	40	227.3 (4-900)	34	269 (23.6-1004.1)	0.3392	38	331.5 (55.5-945)	36	257 (27-1026.4)	0.3687
EDN1	CD68	42	624.1 (45.8-1681.3)	36	470.8 (52.8-2190.6)	0.7688	40	1378.9 (58.1-3861)	36	1019 (107.8-2953.3)	0.0616 (*)
EDN1	CD68intra	42	47 (0.5-225.4)	36	58.7 (7.9-170.3)	0.8303	40	76.8 (0-309.5)	36	51.2 (0-147.7)	0.0718 (*)
EDN1	CD1a	42	3.3 (0-22.5)	37	2.1 (0-8.4)	0.5541	39	4.3 (0-24.3)	32	2.6 (0-13.5)	0.5178
EGFR	CD3	36	339.8 (13-1372)	33	235 (55-858.1)	0.7975	34	510.9 (62.3-1180)	33	679.9 (95.8-1310)	0.2690
EGFR	CD3intra	17	42 (0-126)	12	7 (0-64)	0.2199	15	24 (0-77)	12	14 (0-202)	0.6236
EGFR	T.Bet	16	2.6 (0-9.8)	13	0 (0-3.6)	0.0967 (*)	12	2 (0-13.8)	9	1 (0-14.6)	1.0000
EGFR	GZMB	17	35.6 (0-220.9)	11	14.3 (0-71.3)	0.3429	12	129.5 (12.4-403)	10	49.8 (3.9-219)	0.0300 *
EGFR	CD57	39	31.7 (0-158.8)	38	36.1 (1.6-140.5)	0.7100	35	65.1 (0-281.8)	36	59.4 (0.9-239.5)	0.7620
EGFR	CD8	42	106 (6.3-675)	36	77 (5-377)	0.2835	33	143.1 (18-454)	32	164.7 (3-707.6)	0.9844
EGFR	CD8intra	17	130 (0-232)	13	39.8 (0-116)	0.1369	11	42.3 (0-396)	9	91.7 (0-262)	0.8488
EGFR	CD45RO	38	282.9 (4-1137.1)	36	206.6 (44.8-900)	0.9187	38	322.9 (30-1026.4)	36	257 (27-1026.4)	0.7840
EGFR	CD68	40	613.7 (45.8-1804.4)	38	548.6 (52.8-2165.2)	0.8463	38	1174.2 (58.1-3371)	38	1331.4 (133.2-3012.5)	0.7449
EGFR	CD68intra	40	61.7 (0.5-225.4)	38	50.1 (5-182.8)	0.9802	38	66.2 (0-275.5)	38	50.8 (0-249.2)	0.3261
EGFR	CD1a	42	3.3 (0-19.6)	37	1.9 (0-35.2)	0.6446	36	3.5 (0-23.5)	35	2.6 (0-15.9)	0.8939
ICAM1	CD3	35	349.3 (13-1372)	34	191.6 (14-842)	0.0105 *	35	586 (62.3-1474.6)	32	550.9 (95.8-1310)	0.2362
ICAM1	CD3intra	12	47 (0-156)	17	7 (0-126)	0.0471 *	12	14 (0-298)	15	21 (0-114)	0.7497
ICAM1	T.Bet	12	1.8 (0-9.8)	17	1.8 (0-16.8)	0.9633	7	5.3 (0-131.5)	14	0 (0-14.6)	0.1441
ICAM1	GZMB	12	50.4 (0-167.5)	16	18.2 (0-49.9)	0.2420	9	154 (38.6-280)	13	63.4 (3.9-197)	0.0304 *
ICAM1	CD57	40	43.2 (1.6-139.5)	37	13.7 (0-94.2)	0.0087 **	36	99.9 (4-281.8)	35	39.7 (0-193.3)	0.0134 *
ICAM1	CD8	42	114.1 (5.9-473)	36	55.3 (5-371)	0.0524 (*)	36	206 (18-747)	29	143 (3-471)	0.0827 (*)
ICAM1	CD8intra	13	81 (0-232)	17	39.8 (0-138)	0.3352	9	98.9 (0-242)	11	42.3 (0-394)	0.5935
ICAM1	CD45RO	38	380.2 (14.5-1137.1)	36	137.5 (4-524)	0.0015 **	38	370.5 (39-1152.7)	36	224 (27-773.2)	0.0069 **
ICAM1	CD68	40	689.5 (45.8-1948.6)	38	472.8 (52.8-1908.7)	0.1580	40	1365.5 (195.8-3546.2)	36	1041.9 (58.1-3045.6)	0.1756
ICAM1	CD68intra	40	56.4 (0.5-232.2)	38	54.2 (4.2-164.7)	0.8854	40	54.8 (0-249.2)	36	54.2 (0-249.2)	0.7988
ICAM1	CD1a	41	3.1 (0-19.5)	38	1.8 (0-19.6)	0.9960	38	2.6 (0-24.3)	33	3.9 (0-15.2)	0.6009
MADCAM1	CD3	38	196.3 (13-1372)	31	318 (69.2-789.1)	0.4461	37	559.9 (62.3-1310)	30	566.6 (129.3-1474.6)	0.3902
MADCAM1	CD3intra	19	36 (0-126)	10	10 (0-94)	0.7288	18	19.5 (0-181)	9	17 (0-56)	0.6231
MADCAM1	T.Bet	19	1.8 (0-16.8)	10	1.7 (0-9.7)	0.7569	14	2.9 (0-14.6)	7	0 (0-13.8)	0.3434
MADCAM1	GZMB	18	21.9 (0-99.8)	10	31 (0-167.5)	0.9614	16	79.9 (3.9-280)	6	144.5 (7.5-172)	0.6931
MADCAM1	CD57	39	27 (0-116.5)	38	39.8 (0-158.8)	0.1537	37	65.1 (1.8-268)	34	57.4 (0-281.8)	0.7793
MADCAM1	CD8	40	62.8 (5.9-473)	38	115.1 (5-377)	0.1886	31	158 (7.5-520.5)	34	140 (3-560.9)	0.4767
MADCAM1	CD8intra	19	71.8 (0-178)	11	49 (0-232)	0.8294	12	90.4 (0-242)	8	67 (0-596)	0.9691
MADCAM1	CD45RO	37	156.9 (4-1004.1)	37	292.8 (23.8-793)	0.3660	38	236.5 (27-827.3)	36	404.7 (30-1152.7)	0.0093 **
MADCAM1	CD68	39	458.5 (52.8-1256.5)	39	789.9 (45.8-3022.7)	0.0138 *	38	1222.8 (58.1-2657.6)	38	1358.3 (107.8-3546.2)	0.6236
MADCAM1	CD68intra	39	33.5 (0.5-154)	39	77.1 (3.6-232.2)	0.0112	38	73.9 (0-275.5)	38	45.9 (0-147.7)	0.0923 (*)
MADCAM1	CD1a	40	1.8 (0-19.6)	39	3.1 (0-19.5)	0.8610	39	4.4 (0-24.3)	32	2.3 (0-13.5)	0.2234
VCAM1	CD3	34	355.8 (13-1130)	35	201 (41-657.8)	0.0942 (*)	33	427.1 (62.3-1310)	34	670 (95.8-1474.6)	0.2745
VCAM1	CD3intra	15	10 (0-156)	14	30.5 (0-95)	0.9474	14	18 (0-114)	13	17 (0-77)	0.9611
VCAM1	T.Bet	15	2.2 (0-9.7)	14	0.9 (0-7.1)	0.4149	11	0 (0-14.6)	10	1.5 (0-13.8)	0.8522
VCAM1	GZMB	15	40 (0-220.9)	13	14.3 (0-71.3)	0.1707	12	100.7 (7.5-280)	10	82.7 (3.9-219)	0.7223
VCAM1	CD57	39	37.8 (1.7-158.8)	38	26.9 (0-140.5)	0.1582	34	63.2 (0.9-195.2)	37	60.7 (0-317.5)	0.6846
VCAM1	CD8	41	134 (5-675)	37	57.9 (5.9-372.9)	0.0686 (*)	35	171.4 (3-747)	30	144.5 (7.5-520.5)	0.5174
VCAM1	CD8intra	16	49.4 (0-509)	14	56.7 (0-178)	0.8678	11	81.9 (0-242)	9	98.9 (0-394)	0.5674
VCAM1	CD45RO	39	292.8 (4-1137.1)	35	156.9 (23.8-786.5)	0.4379	38	347.9 (30-1026.4)	36	245.8 (27-716.7)	0.2556
VCAM1	CD68	40	750.6 (58.3-3022.7)	38	419.1 (45.8-1631.3)	0.0084	38	1564.2 (58.1-4187.3)	38	1046.3 (133.2-2657.6)	0.0697 (*)
VCAM1	CD68intra	40	73.7 (0.5-261.9)	38	35.7 (3.6-164.7)	0.0385 *	38	43.4 (0-221.8)	38	66.2 (0-249	

position	Parameters	Cluster	position	Parameters	Cluster	position	Parameters	Cluster	position	Parameters	Cluster
1	Qty_tot_Vb6.1L09	A	44	Qty_tot_Vb6.5L04	C	87	Qty_tot_Vb7L05	C	130	Qty_tot_Vb3L06	E
2	Qty_tot_Vb6.5L09	A	45	Qty_tot_Vb5.2L06	C	88	Qty_tot_Vb6.4L06	C	131	Qty_tot_Vb3L08	E
3	Qty_tot_Vb5.2L05	A	46	Qty_tot_Vb6.1L05	C	89	Qty_tot_Vb18L06	C	132	E-dCT-CXCL10	E
4	Qty_tot_Vb24L08	A	47	Qty_tot_Vb6.5L05	C	90	Qty_tot_Vb16L03	C	133	Qty_tot_Vb2L03	E
5	Qty_tot_Vb23L04	A	48	Qty_tot_Vb6.1L03	C	91	Qty_tot_Vb5.2L07	C	134	Qty_tot_Vb11L05	E
6	Qty_tot_Vb23L07	A	49	Qty_tot_Vb3L05	C	92	Qty_tot_Vb9L03	C	135	Qty_tot_Vb5.1L08	E
7	Qty_tot_Vb5.2L09	A	50	Qty_tot_Vb18L05	C	93	Qty_tot_Vb16L08	C	136	E-dCT-CXCR6	E
8	Qty_tot_Vb23L10	A	51	Qty_tot_Vb6.4L08	C	94	Qty_tot_Vb6.5L03	C	137	Qty_tot_Vb8L05	E
9	Qty_tot_Vb22L09	A	52	Qty_tot_Vb17L07	C	95	Qty_tot_Vb8L07	C	138	Qty_tot_Vb8L08	E
10	Qty_tot_Vb21L08	A	53	Qty_tot_Vb22L08	C	96	Qty_tot_Vb2L07	C	139	Qty_tot_Vb11L09	E
11	Qty_tot_Vb1L03	A	54	Qty_tot_Vb17L09	C	97	Qty_tot_Vb7L10	C	140	Qty_tot_Vb2L10	E
12	Qty_tot_Vb4L05	B	55	Qty_tot_Vb11L06	C	98	Qty_tot_Vb11L04	C	141	E-dCT-CX3CL1	E
13	Qty_tot_Vb24L04	B	56	Qty_tot_Vb11L08	C	99	Qty_tot_Vb7L07	C	142	Qty_tot_Vb4L04	E
14	Qty_tot_Vb23L09	B	57	Qty_tot_Vb2L09	C	100	Qty_tot_Vb21L03	C	143	Qty_tot_Vb24L06	E
15	Qty_tot_Vb22L10	B	58	Qty_tot_Vb2L06	C	101	Qty_tot_Vb1L10	C	144	E-dCT-CXCL9	E
16	Qty_tot_Vb24L10	B	59	Qty_tot_Vb24L07	C	102	Qty_tot_Vb18L09	C	145	Qty_tot_Vb17L03	E
17	Qty_tot_Vb4L09	B	60	Qty_tot_Vb6.1L08	C	103	Qty_tot_Vb16L10	C	146	Qty_tot_Vb23L08	E
18	Qty_tot_Vb5.2L10	B	61	Qty_tot_Vb6.5L08	C	104	Qty_tot_Vb15L05	C	147	Qty_tot_Vb18L10	E
19	Qty_tot_Vb15L03	B	62	Qty_tot_Vb9L08	C	105	Qty_tot_Vb2L04	D	148	Qty_tot_Vb17L10	E
20	Qty_tot_Vb5.1L10	B	63	Qty_tot_Vb6.4L04	C	106	Qty_tot_Vb4L07	D	149	Qty_tot_Vb5.1L07	E
21	Qty_tot_Vb24L09	B	64	Qty_tot_Vb22L07	C	107	Qty_tot_Vb8L06	D	150	Qty_tot_Vb5.2L03	E
22	Qty_tot_Vb1L09	B	65	Qty_tot_Vb18L08	C	108	Qty_tot_Vb7L09	D	151	Qty_tot_Vb16L04	E
23	Qty_tot_Vb16L09	B	66	Qty_tot_Vb2L05	C	109	Qty_tot_Vb9L06	D	152	Qty_tot_Vb6.4L07	E
24	Qty_tot_Vb15L04	B	67	Qty_tot_Vb3L09	C	110	Qty_tot_Vb8L10	D	153	Qty_tot_Vb8L09	E
25	Qty_tot_Vb3L07	C	68	Qty_tot_Vb4L08	C	111	Qty_tot_Vb17L04	D	154	Qty_tot_Vb5.2L08	E
26	Qty_tot_Vb6.4L03	C	69	Qty_tot_Vb16L05	C	112	Qty_tot_Vb4L10	D	155	Qty_tot_Vb17L06	E
27	Qty_tot_Vb6.1L06	C	70	Qty_tot_Vb3L04	C	113	Qty_tot_Vb9L05	D	156	Qty_tot_Vb1L04	E
28	Qty_tot_Vb6.5L06	C	71	Qty_tot_Vb6.1L07	C	114	Qty_tot_Vb5.1L09	D	157	Qty_tot_Vb11L10	E
29	Qty_tot_Vb11L07	C	72	Qty_tot_Vb6.5L07	C	115	Qty_tot_Vb22L05	D	158	Qty_tot_Vb3L10	E
30	Qty_tot_Vb22L06	C	73	Qty_tot_Vb5.1L04	C	116	Qty_tot_Vb1L05	D	159	Qty_tot_Vb18L03	E
31	Qty_tot_Vb18L04	C	74	Qty_tot_Vb7L06	C	117	Qty_tot_Vb15L07	D	160	Qty_tot_Vb17L08	E
32	Qty_tot_Vb9L04	C	75	Qty_tot_Vb7L08	C	118	Qty_tot_Vb21L06	D	161	Qty_tot_Vb21L05	E
33	Qty_tot_Vb11L03	C	76	Qty_tot_Vb6.4L05	C	119	Qty_tot_Vb7L03	D	162	Qty_tot_Vb15L06	E
34	Qty_tot_Vb4L06	C	77	Qty_tot_Vb9L07	C	120	Qty_tot_Vb1L06	D	163	Qty_tot_Vb9L10	F
35	Qty_tot_Vb18L07	C	78	Qty_tot_Vb21L09	C	121	Qty_tot_Vb21L10	D	164	Qty_tot_Vb17L05	F
36	Qty_tot_Vb1L08	C	79	Qty_tot_Vb22L04	C	122	Qty_tot_Vb6.4L09	D	165	Qty_tot_Vb24L03	F
37	Qty_tot_Vb5.1L05	C	80	Qty_tot_Vb6.4L10	C	123	Qty_tot_Vb22L03	D	166	Qty_tot_Vb6.5L10	F
38	Qty_tot_Vb5.2L04	C	81	Qty_tot_Vb21L04	C	124	Qty_tot_Vb16L06	D	167	Qty_tot_Vb23L03	F
39	Qty_tot_Vb2L08	C	82	Qty_tot_Vb16L07	C	125	Qty_tot_Vb15L08	D	168	Qty_tot_Vb23L05	F
40	Qty_tot_Vb8L04	C	83	Qty_tot_Vb7L04	C	126	E-dCT-CXCR3	E	169	Qty_tot_Vb6.1L10	F
41	Qty_tot_Vb5.1L03	C	84	Qty_tot_Vb9L09	C	127	Qty_tot_Vb24L05	E	170	Qty_tot_Vb15L09	F
42	Qty_tot_Vb5.1L06	C	85	Qty_tot_Vb1L07	C	128	Qty_tot_Vb8L03	E	171	Qty_tot_Vb3L03	F
43	Qty_tot_Vb6.1L04	C	86	Qty_tot_Vb21L07	C	129	Qty_tot_Vb23L06	E	172	Qty_tot_Vb15L10	F

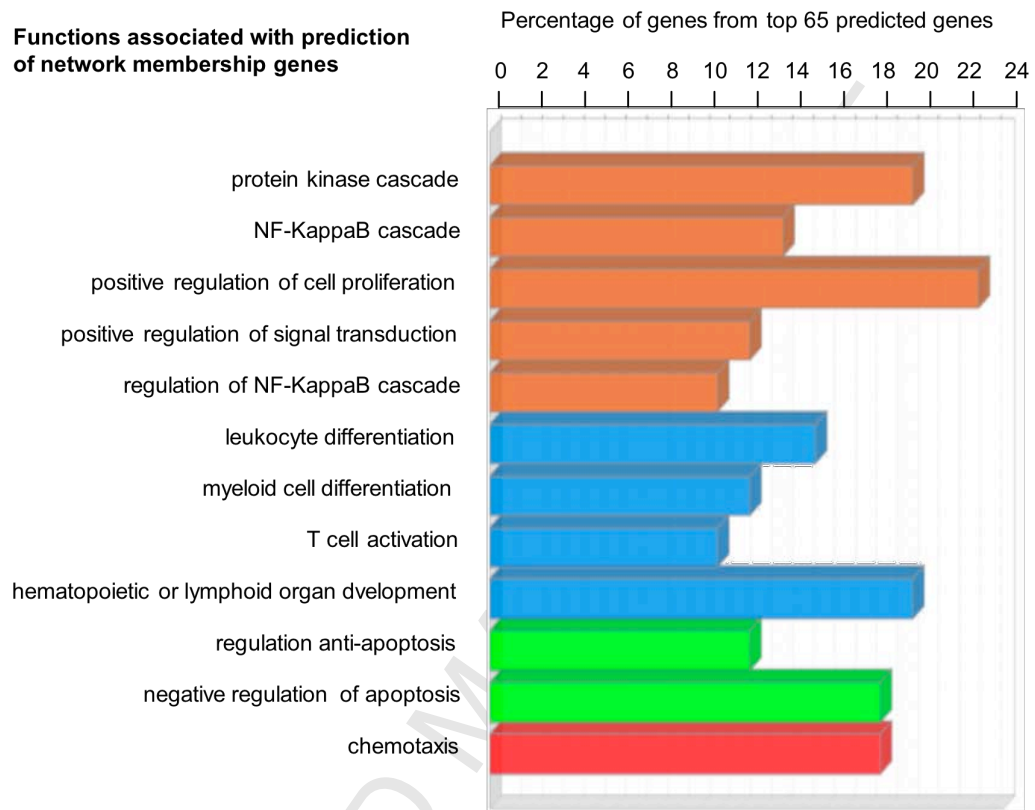
Supplementary Table 5



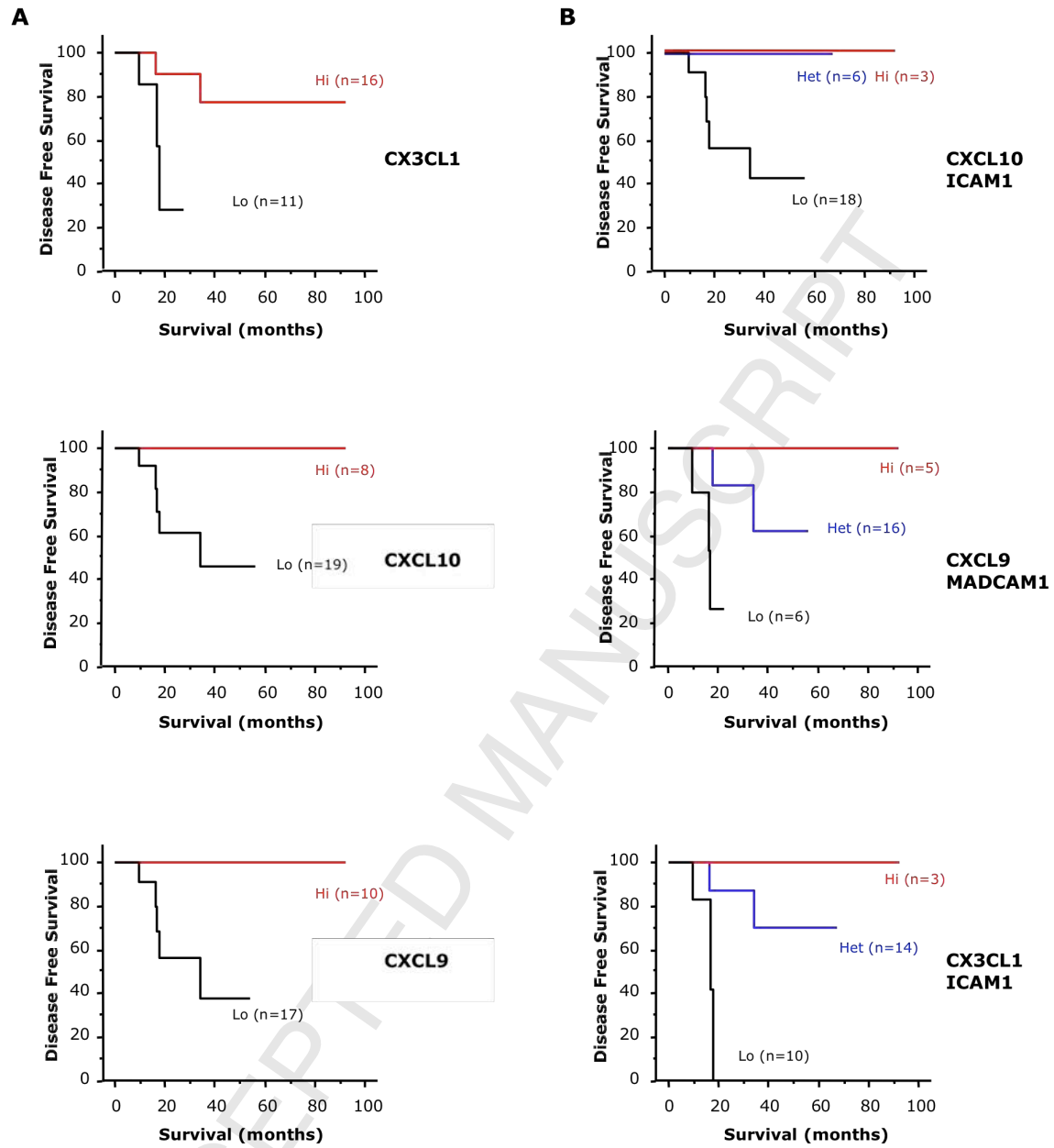
Cohorts	Percentage (total number) of patients				pv*
	TMA	Taqman	FACS	TCR	
<b>Tumor (T) stage †</b>					0.8635
pT1-2	12.4 (13)	12.4 (13)	15.4 (4)	10 (1)	
pT3	72.4 (76)	72.4 (76)	73.1 (19)	80 (8)	
pT4	15.2 (16)	15.2 (16)	11.5 (3)	10 (1)	
<b>N stage</b>					NA
N0	47.6 (50)	47.6 (50)	46.2 (12)	50 (5)	
N+	52.4 (55)	52.4 (55)	53.8 (14)	50 (5)	
<b>M stage</b>					
M0	67.6 (71)	67.6 (71)	73.1 (19)	70 (7)	
M+	32.4 (34)	32.4 (34)	26.9 (7)	30 (3)	
<b>Age (yrs)</b>					0.9955
< 65	26.7 (28)	26.7 (28)	30.8 (8)	30 (3)	
65-75	23.8 (25)	23.8 (25)	15.4 (4)	40 (4)	
75-85	38.1 (40)	38.1 (40)	38.5 (10)	20 (2)	
> 85	11.4 (12)	11.4 (12)	15.4 (4)	10 (1)	
<b>Gender</b>					NA
Male	55.2 (58)	55.2 (58)	61.5 (16)	50 (5)	
Female	44.8 (47)	44.8 (47)	38.5 (10)	50 (5)	
<b>Tumor Location</b>					0.9810
Right colon	26.7 (28)	26.7 (28)	26.9 (7)	30 (3)	
Left colon	12.4 (13)	12.4 (13)	7.7 (2)	0 (0)	
Sigmoid colon	44.8 (47)	44.8 (47)	38.5 (10)	50 (5)	
Rectum	16.2 (17)	16.2 (17)	26.9 (7)	20 (2)	
<b>Number of lymph nodes</b>					NA
<8	30.5 (32)	30.5 (32)	7.7 (2)	30 (3)	
>=8	69.5 (73)	69.5 (73)	92.3 (24)	70 (7)	
<b>Histological grade</b>					1.0000
Well Differentiated	67.6 (71)	67.6 (71)	46.2 (12)	50 (5)	
Moderately Differentiated	28.6 (30)	28.6 (30)	50 (13)	40 (4)	
Poorly Differentiated	2.9 (3)	2.9 (3)	0 (0)	10 (1)	
NaN	1 (1)	1 (1)	3.8 (1)	0 (0)	
<b>Mucinous Colloid Type</b>					0.9898
No	75.2 (79)	75.2 (79)	84.6 (22)	70 (7)	
Yes	21.9 (23)	21.9 (23)	11.5 (3)	30 (3)	
NaN	2.9 (3)	2.9 (3)	3.8 (1)	0 (0)	
<b>Occlusion</b>					NA
No	86.7 (91)	86.7 (91)	96.2 (25)	90 (9)	
Yes	13.3 (14)	13.3 (14)	3.8 (1)	10 (1)	
<b>Perforation</b>					NA
No	97.1 (102)	97.1 (102)	100 (26)	100 (10)	
Yes	2.9 (3)	2.9 (3)	0 (0)	0 (0)	
<b>Pre Operation Treatment</b>					
No	98.1 (103)	98.1 (103)	100 (26)	100 (10)	
Yes	1.9 (2)	1.9 (2)	0 (0)	0 (0)	
<b>Post Operation Treatment</b>					NA
No	58.1 (61)	58.1 (61)	61.5 (16)	50 (5)	
Yes	41.9 (44)	41.9 (44)	38.5 (10)	50 (5)	
<b>VELIPI</b>					0.9898
No	63.8 (67)	63.8 (67)	57.7 (15)	70 (7)	
Yes	34.3 (36)	34.3 (36)	34.6 (9)	20 (2)	
NaN	1.9 (2)	1.9 (2)	7.7 (2)	10 (1)	

† The stage was determined by pathological (p) examination. Tis denotes carcinoma in situ, T1 tumor invading submucosa, T2 tumor invading muscularis propria, T3 tumor penetrating muscularis propria and invading subserosa, and T4 tumor completely through the intestine's wall and invading serosa and nearby tissues. #: Bowel perforation presentation. pv\* kruskal-Wallis P value. NA: Not applicable.

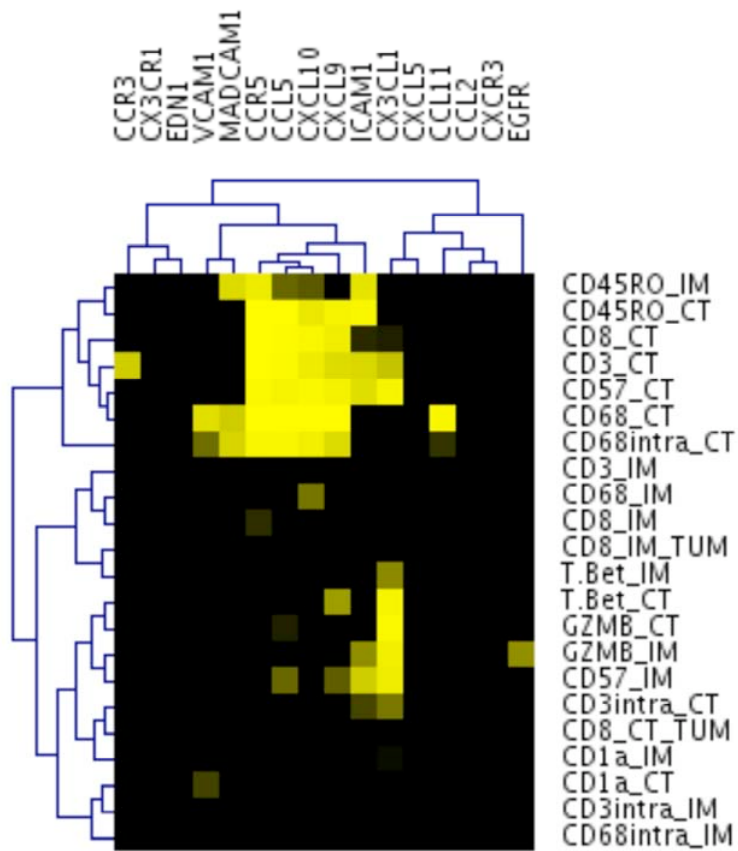
Supplementary Table 6



Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3

## Supplementary experimental procedures

### Patients and database

The records of colorectal cancer (CRC) patients who underwent a primary resection of their tumor at the Laennec-HEGP Hospitals between 1996 and 2004 were reviewed and previously described (Galon et al. 2006). Frozen tumor samples available from Laennec-HEGP Hospitals from 1996-2004, with sufficient RNA quality and quantity, were selected (validation cohort 1, n=108; and re-validation independent cohort 2, n=27). The RNA samples analyzed were from 108 and 27 different patients. These patients were used for gene expression experiments (Taqman cohort), and for tissue-microarrays experiments (TMA cohort). The intratumoral T cell repertoire was performed on 10 randomly selected colorectal tumors from Laennec-HEGP Hospitals from 1996-2004, for which sufficient RNA quality and quantity were available (TCR cohort). All patients from Laennec-HEGP Hospitals from 2002-2004, from the cohort of 108 patients, and from which fresh tumors were available, were used for flow cytometry experiment (FACS cohort) (see Table S6 for details). The observation time in the cohorts was the interval between diagnosis and last contact (death or last follow-up). Data were censored at the last follow-up for patients without relapse, or death. The min:max values until progression/death or last follow-up were (0:136) months, respectively. Three patients for whom follow-up data were unavailable were excluded from survival analysis. Time to recurrence or disease-free time was defined as the interval from the date of surgery to confirmed tumor relapse date for relapsed patients and from the date of surgery to the date of last follow-up for disease-free patients.

Histopathological and clinical findings were scored according to the UICC-TNM staging system<sup>1,2</sup>. Post-surgical patient surveillance was performed at Laennec-

HEGP Hospitals for all patients according to general practice for CRC patients. Adjuvant chemotherapy was administered to patients with stage III CRCs, to high-risk stage II CRCs, and palliative chemotherapy to patients with advanced colorectal cancers (stage IV) and to patients without complete resection of the tumor. Adjuvant chemotherapy was fluorouracil (FU)-based. Follow-up data were collected prospectively and updated. A secure Web-based database, TME.db (Tumor MicroEnvironment Database), was built on a 3-tier architecture using Java-2 Enterprise-Edition (J2EE) to integrate the clinical data and the data from high-throughput technologies. Table S5 summarizes the clinical characteristics of the patients. The 5 cohorts of patients used had similar clinical characteristics with no statistical difference between the cohorts (all P value > 0.86).

### **Gene expression analysis**

Tissue samples were snap-frozen within 15 minutes after surgery and stored in liquid nitrogen. Frozen tumor samples (cohort 1, n=108; and re-validation independent cohort 2, n=27) of randomly selected patients available from Laennec-HEGP Hospitals (1996-2004), with sufficient RNA quality and quantity, were selected for gene expression analysis. The RNA samples analyzed were from 108 and 27 different patients. Total RNA was isolated by homogenization with RNeasy isolation-kit (Qiagen, Valencia, CA). The integrity and the quantity of the RNA were evaluated on a bioanalyzer-2100 (Agilent Technologies, Palo Alto, CA). RT-PCR experiments were performed according to the manufacturer's instructions (Applied-Biosystems, Foster City, CA). Quantitative real-time TaqMan-PCR was performed using Low-Density-Arrays and the 7900 robotic real-time PCR-system (Applied-Biosystems). 18S ribosomal RNA primers and probe were used as internal control. Gene expression analyses were performed using Ct-values (threshold cycle) normalized to

18S ribosomal RNA ( $\Delta$ Ct). Data were analyzed using the SDS Software v2.2 (Applied-Biosystems) and TME statistical module.

### **Large-scale flow-cytometric analysis**

After mechanical dispersion, cells from 39 fresh tumor samples were washed and subjected to four-color flow-cytometry. Cells were resuspended in PBS/0.5% BSA and incubated for 30 minutes at 4°C with antibodies for large-scale phenotypic analysis of T cells, and with relevant isotype controls. 40,000 cells were analyzed per run. Analyses were performed with a FACScalibur flow-cytometer and CellQuest software (Becton Dickinson, San Diego, CA). Parameters (n=149) were measured as percentage of total cells and percentage of cells within CD3+ cells. Spearman hierarchical clustering was applied to the correlation matrix and the results were displayed by using the Genesis program (software available at <http://www.genome.tugraz.at>).

### **T cell receptor repertoire analysis**

The CDR3 length distribution analysis method is a molecular biology tool based on the analysis of V and V-J gene composition and transcript lengths. Following reverse-transcription into cDNA, CDR3-encoding mRNA is amplified by PCR using specific V and C primers. PCR products are then labeled with a fluorescent C or J primer. Finally, amplified fragments are separated with electrophoresis, and visualized using fluorescent sequencing equipment. Bands are separated by a space corresponding to a transcript length of three nucleotides; suggesting that the population is composed of in-frame transcripts of different length. Signal processing programs are used to digitize the gel image thus enabling quantification of the relative amount of fluorescence in each band. A frequency

distribution of the CDR3 lengths is then obtained. Once the relative proportions of each length class of mRNA are calculated, the CDR3 length distribution can be displayed as a probability distribution. In humans, the TCR chain has been described as having 65 V genes grouped into 30 families on the basis of the sequence similarity<sup>3</sup>. Among these different V genes families, we used the 26 families that have been shown to lead to the expression of a functional TCR (table below).

The intratumoral T cell repertoire was performed on 10 randomly selected colorectal tumors using the TcLandscape technology (TcLand, Nantes, France). The quantitative PCR assessment for each of the 26 TCR V $\beta$  chain families was performed. The data were expressed in ratio V $\beta$ /HPRT and their quantities were represented by the height of the peak in the TcLandscapes and on the histograms. T cell repertoire analysis was performed by combining qualitative alterations of V $\beta$  usage at the CDR3 length level (13 different CDR3 length) with the magnitude of expression of each V $\beta$  mRNA species. The qualitative data are composed of the Immunoscope profiles. The analysis of these profiles consists in computing a percentage of alteration for each V $\beta$  chain families<sup>3, 4</sup>.

TcLandscape V $\beta$ families	Analyzed genes*
Vb 1	1S1
Vb 2	2S1
Vb 3	3S1
Vb 4	4S1
Vb 5.1	5S1
Vb 5.2	5S2, 5S3, 5S4, 5S6
Vb 6.1	6S1, 6S2, 6S3, 6S6, 6S8
Vb 6.4	6S4
Vb 6.5	6S5



Vb 7	7S1, 7S2, 7S3
Vb 8	8S1, 8S2, 8S3
Vb 9	9S1
Vb 11	11S1
Vb 12.1	12S1
Vb 12.2	12S2, 12S3
Vb 13.1	13S1, 13S2, 13S3, 13S6, 13S9
Vb 13.5	13S5
Vb 14	14S1
Vb 15	15S1
Vb 16	16S1
Vb 17	17S1
Vb 18	18S1
Vb 21	21S1, 21S2, 21S3
Vb 22	22S1
Vb 23	23S1
Vb 24	24S1

\* V $\beta$  subfamilies according to the nomenclature of Arden<sup>5</sup>.

### Tissue Microarray Construction

Using a tissue-array instrument (Beecher Instruments, ALPHELYS, Plaisir, France), two representative regions of the tumor (center (CT) and invasive margin (IM)) were punched (0.6 mm and 1 mm-diameter, respectively) from paraffin-embedded tissue-blocks. CRCs (n=108) that were resected at Laennec-HEGP Hospitals between 1990 and 2004 were used for construction of tissue microarrays, and four cores per tumor were spotted (two for each tumor region). Based on pathologic findings, tumor spreading (T-stage), lymph node invasion (N-stage), and metastasis (M-stage), patients with these tumors were representative for the entire cohort. Tissue-cores arrayed into recipient paraffin-blocks were cut into 5  $\mu$ m sections for Harris's hematoxylin (HE) and immunohistochemical stainings.

## Immunohistochemistry

Tissue microarray sections were incubated (60 min. at room temperature) with monoclonal antibodies against CD3 (SP7), CD8 (4B11), CD45RO (OPD4), GZMB (GrB-7), CD57 (NK1), CD1A (O10), cytokeratin (AE1AE3) and cytokeratin-8 (Neomarkers, Fremont, CA), T-bet (4B10) (Santa Cruz Biotechnology, Santa Cruz, CA), and CD68 (PG-M1) (Dako, Copenhagen, Denmark). Envision+ system (enzyme-conjugated polymer backbone coupled to secondary antibodies) and DAB-chromogen were applied (Dako). Double stainings were revealed with phosphate-conjugated secondary antibodies and FastBlue-chromogen. For single stainings, tissue sections were counterstained with Harris' hematoxylin. Isotype-matched mouse monoclonal and rabbit polyclonal antibodies were used as negative controls. Slides were analyzed using an image analysis workstation (Spot Browser®, ALPHELYS). Polychromatic high-resolution spot-images (740x540 pixel, 1.181 mm/pixel resolution) were obtained (x100 fold magnification). The density was recorded as the number of positive cells per unit tissue surface area. In each tumor region (CT and IM), and for each duplicate, the mean density was used for statistical analysis.

## Statistical analysis and network reconstruction

The gene prediction network in Fig. 1 was created using the STRING database v6.3<sup>6</sup>. The prediction with STRING was performed based on a subset of 12 genes which showed a significant log-rank p-value for DFS at the median cut-off level for high and low expression. The gene expression data was acquired by a RT-PCR study for 47 immune-related genes in a cohort of 108 colorectal cancer (CRC) patients (table S1). As first step the “top 65” genes were predicted by STRING. Three

among those genes (CD4, PTGS2, and VEGF) were re-predicted as already part of the 47 RT-PCR genes. CX3CL1 was the first one predicted. The gene network was created based on the *in silico* predicted genes and the 47 RT-PCR genes from experimental data. The node sizes of the network are based on the Hazard Ratio for DFS, determined by using the median cut-off approach (table S1). The nodes surrounded by a black border are the genes significantly (log-rank) associated with DFS. The nodes surrounded by a red border are predicted by STRING (the thick border for the “top 10” predicted ones). All other nodes are colored based on their single or multiple occurrence in different GO categories. The edge weights of the network are based on the integrated score of the pairwise un-centered Pearson correlation value between the 49 RT-PCR genes and the combined edge scores for all genes provided by STRING. To integrate the correlation and the scores obtained from STRING the method for calculating the combined scores<sup>6</sup> was applied. The combined score for the edges generated by STRING is based on the sum of following relations: Neighborhood, Gene Fusion, Cooccurrence, Coexpression, Experiments, Databases, Textmining and Homology.

The visualization of the network was done using the open-source Cytoscape<sup>7</sup> software platform for visualizing and integrating biomolecular interaction networks. To layout the network, the Golorize<sup>8</sup> plug-in for the network node visualization was used. The layout algorithm determines the node positions based on Gene Ontology (GO) categories of the network genes<sup>9</sup>. The implementation of Golorize is based on the BiNGO<sup>10</sup> plug-in. Both plug-ins are used within Cytoscape. Edge thickness levels on the figure show the relation strength based on the integrated score value between the nodes. Dashed lines:  $R < 0.65$ , thin solid lines:  $0.65 \leq R < 0.8$  and thick solid lines:  $R \geq 0.8$ . Gray lines indicate the STRING origin of the edge relation, blue lines

indicate the un-centered Pearson correlation edge relation and pink lines indicate the overlap of *in silico* and experimental correlation relations.

Genesis clustering software was used to visualize the correlation matrix and hazard ratios matrix using Pearson un-centered hierarchical clustering with average linkage. Kaplan Meier estimators of survival were used to visualize the survival curves. The log-rank test was used to compare disease-free and overall survival between patients in different groups. For initial network reconstruction, markers were measured with RT-PCR and the median expression level was taken as cut-off to dichotomize the variables. For experimental validation of the predicted genes, the “minimum p-value” cut-off for each marker was calculated. Hazard ratios were corrected using the formula proposed by Hollander<sup>12</sup>. Hazard ratio-matrix (heatmap) followed by unsupervised hierarchical clustering was represented from favorable prognosis HR=0.4 (red) to poor prognosis HR=2.5 (blue). Hazard ratios were calculated for high and low gene expression compared to the whole cohort. All combinations of hazard ratios were uploaded into the Genesis clustering program. Each cell on the matrix represents the hazard ratio obtained by a pair of genes when a Cox proportional hazards analysis is applied to the combined high/low expression data for these genes (the diagonal represents an analysis based on just the high/low data for that gene). P-values for the HiHi vs. LoLo gene combination analysis were corrected for multiple testing using the Benjamini-Hochberg method. All HR with  $HR < 0.55$  or  $HR > 1.66$ , were significant.

We applied the Kruskal-Wallis one-way analysis of variance to determine if any of the patient cohorts was significant different regarding the clinical parameters. No significant difference was found between cohorts. All through the text a p-value <

0.05 was considered statistically significant. All analyzes were performed with the statistical software R (survival package) and Statview.

## Supplementary References

1. Dukes C. The classification of cancer of the rectum. *J Pathol* 1932;35:323.
2. Weitz J, Koch M, Debus J, Hohler T, Galle PR, Buchler MW. Colorectal cancer. *Lancet* 2005;365:153-65.
3. Folch G, Lefranc MP. The human T cell receptor beta variable (TRBV) genes. *Exp Clin Immunogenet* 2000;17:42-54.
4. Miqueu P, Guillet M, Degauque N, Dore JC, Soullillou JP, Brouard S. Statistical analysis of CDR3 length distributions for the assessment of T and B cell repertoire biases. *Mol Immunol* 2007;44:1057-64.
5. Arden B, Clark SP, Kabelitz D, Mak TW. Human T-cell receptor variable gene segment families. *Immunogenetics* 1995;42:455-500.
6. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* 2005;33:D433-7.
7. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498-504.
8. Garcia O, Saveanu C, Cline M, Fromont-Racine M, Jacquier A, Schwikowski B, Aittokallio T. GOLORize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. *Bioinformatics* 2007;23:394-6.
9. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;25:25-9.
10. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 2005;21:3448-9.
11. Werhli AV, Grzegorzczak M, Husmeier D. Comparative evaluation of reverse engineering gene regulatory networks with relevance networks, graphical gaussian models and bayesian networks. *Bioinformatics* 2006;22:2523-31.

12. Hollander N, Sauerbrei W, Schumacher M. Confidence intervals for the effect of a prognostic factor after selection of an 'optimal' cutpoint. Stat Med 2004;23:1701-13.

ACCEPTED MANUSCRIPT

## **Article 4**

"Analysis of infiltrating cytotoxic, Th1, Th2, Treg and Th17 cells in patients with colorectal cancer : impact on clinical outcome"

Marie Tosolini, Bernhard Mlecnik, Stéphanie Mauger, Tessa Fredriksen, Amos Kirilovsky, Gabriela Bindea, Anne Berger, Tchao Meatchi, Zlatko Trajanoski, Wolf-Herman Fridman, Franck Pagès, Jérôme Galon

en cours d'écriture

Nous avons précédemment démontré l'importance de la réponse immunitaire avec une coordination efficace dans le contrôle de l'invasion de ganglions ou d'organes à distance (article 2 [141]). Nous avons également montré l'importance de la réponse immunitaire de type Th1 et mémoire pour le pronostic des patients. Nous avons approfondi ce travail en analysant la réponse cytotoxique en relation avec les sous-populations T helper (Th1, Th2, Th17 et T régulateur).

J'ai effectué les marquages des lymphocytes T CD8 (CD8), les lymphocytes T régulateurs (FoxP3) des lymphocytes Th1 (T-bet) et des lymphocytes Th17 (IL-17) par immunohistochimie sur les Tissue MicroArrays de 110 patients. J'ai analysé ces différents marquages à l'aide du logiciel Spotbrowser au centre et front de progression de la tumeur (figures 7, 8 et figures supplémentaires 3, 4, 5). J'ai effectué des doubles marquages RORC ou IL17 avec CCL24 (figure supplémentaire 1). J'ai également effectué les extractions d'ARN puis les PCR en temps réel sur Low-Density-Array pour l'expression des marqueurs des différentes sous-populations T (figures 1, 2, 3, 4, 5 et figure supplémentaire 2). A partir de ces données, j'ai fait les tests statistiques appropriés, tracé des matrices de corrélations et courbes de Kaplan-Meier.

## Résumé

Le type, la densité et la localisation des cellules immunitaires ont un facteur pronostic supérieur et indépendant aux critères d'extension de la tumeur. Le but de cette étude est d'analyser la balance entre les cellules cytotoxiques et les T helper ainsi que leurs impacts sur la survie sans récurrence des patients. Dans ce but, j'ai analysé l'expression de gènes immuns de 126 tumeurs congelées par PCR quantitative en temps réel. J'ai également quantifié les cellules infiltrantes Th1, Th17, T régulatrices et cytotoxiques par immunohistochimie sur TMA sur deux zones de la tumeur : le centre et le front de progression.

La matrice de corrélation clustérisée de manière non supervisée révèle des clusters fonctionnels de gènes. Nous avons identifié

- 1 cluster composé de gènes Th1 et cytotoxique (IRF1, GZMB, IL27, GNLY, PRF1, CCL5, CD8A, STAT1)
- 1 cluster composé essentiellement de gènes Th1 (IL12RB1, CD28, CCR5, HLA DMB, IL12RB2, CD38, CXCR6, TBX21)
- 3 clusters composés essentiellement de gènes Th2 (IFNGR1, STAT3, IFNGR2, IL10RB, IL4R, STAT6 ; IL4, IL5, IL13 et CCR7, CD3E, CD40LG, CCL19, CCR4, GATA3)
- 1 cluster composé essentiellement de gènes Treg (IL10, TGF $\beta$ )
- 1 cluster composé de gènes Treg et Th2 (FOXP3, CTLA4, CCL17, CCL22)
- 1 cluster composé essentiellement de gènes Th17 (RORC, IL17A, CCL24)

L'impact de ces différents clusters sur le devenir clinique des patients a ensuite été analysé. A cette fin, des groupes de patients ont été définis grâce à l'arbre hiérarchique non supervisé de l'expression des gènes du cluster. Aucun pronostic particulier n'a été mis en évidence pour les clusters Th2 ni pour les clusters Treg. En revanche, le groupe de patients avec une forte expression des gènes Th1 et Th1/cytotoxique présente un taux de survie sans récurrence augmenté par rapport aux patients avec une faible expression. Au contraire, les patients avec une faible expression des gènes Th17 semblent être de bons pronostics. L'analyse combinée des Th1/cytotoxiques et des Th17 permet d'avoir une meilleure discrimination de la récurrence. Aucun patient Th1/cytotoxique fort et Th17 faible ne récidive. A l'opposé, la combinaison la pire en terme de récurrence est Th1/cytotoxique faible et Th17 fort. Ces résultats ont été confirmés par une analyse des protéines *in situ* par immunohistochimie. Les



patients ayant une forte densité de cellules CD8+ (marqueur Th1) dans les deux zones de la tumeur, quelque soit la densité de cellules IL17+ (marqueur Th17) n'ont pas fait de récurrence. A l'opposé, les patients avec une faible densité de cellules CD8+ dans les deux zones de la tumeur sont de mauvais pronostic. Les patients avec une densité non concordante dans les deux zones de la tumeur (hétérogènes), peuvent être séparés grâce à la densité de cellules IL17+. Aucun patient CD8 hétérogène et IL17 faible ne récidive tandis que 80% des patients CD8 hétérogènes IL17 fort récidivent dans les deux ans.

Les clusters fonctionnels associés avec les lymphocytes Th1/ cytotoxique et Th17 ont des effets opposés sur la survie des patients. Leurs informations sont complémentaires bien que l'effet des Th1/cytotoxiques soit supérieur à celles des Th17 sur le pronostic des patients.

## Abstract

*Objectives:* Tumor microenvironment is complex and composed of many immune T cell subpopulations. The aim of the study was to analyze the balance between the cytotoxic and the different subset of helper T-cells in colorectal cancer and the impact on disease-free survival.

*Methods:* The tumor microenvironment was investigated in 126 frozen colorectal tumor samples by analyzing the expression of immune-related genes by Low-Density-Array on real-time PCR Taqman 7900 HT. Infiltrating cytotoxic T cells, Treg, Th1, Th17 cells of colorectal cancer patients were quantified by immunohistochemical analyses of tissue microarrays containing tissue cores from the center and from the invasive margin of the tumor. For pairwise comparisons of parametric and non-parametric data and for survival analysis, the Student's t-test, Wilcoxon rank-sum test and Logrank test were used, respectively.

*Results:* We first investigated the expression of cytotoxic, Th1, Th2, Treg and Th17 genes. Hierarchical clustering of a correlation matrix revealed functional clusters of genes associated with Th17 (RORC, IL17A), Th2 (IL4, IL5, IL13), Th1 (Tbet, IRF1, IL12Rb2, STAT4), and cytotoxicity (GNLY, GZMB, PERF1). High-IL17A mRNA expression level was most frequent at early stages of tumor progression. Patients with high expression of the Th17 cluster had a poor prognosis whereas patients with high expression of the Th1 cluster had prolonged disease-free survival. In contrast no prediction of the prognosis was associated with the Th2 clusters. The combined analysis of cytotoxic/Th1 and Th17 clusters gave a better discrimination for relapse. In situ analysis of IL17+ cells and CD8+ cells using tissue-microarray confirmed with these results.

*Conclusion:* Functional clusters associated with Th1 and Th17 cells have opposite effect on patients survival and bring complementary information.

## Introduction

Cancer progression is a complex process involving host-tumor interactions through multiple molecular and cellular factors of the tumor microenvironment. In mice, immune cells appear to prevent the development of tumors and inhibits tumor progression. However, through inflammation-dependant mechanisms, the innate immune system can promote tumor development [1] [2]. In human, lymphocytes are shown to participate in antitumoral responses. Tumor-infiltrating T cells are associated with improved clinical outcome and survival in colorectal cancer patients, according to findings in melanoma and ovarian cancer. We showed that a high intratumoral memory T cells density correlated with the decrease of early metastatic events and the prevention of relapse in colorectal cancer patients [3]. Furthermore, we demonstrated that the functional orientation, density and location of immune cells profoundly influence the clinical outcome for the patients, regardless of cancer stages [4]. In opposition to patients with metastasis or low density of immune cells, the immune response was coordinated in patients without metastasis and with high density of immune cells [5]. We previously analyzed adaptive immune response, immunosuppression and inflammation. Adaptive immune response correlated with the absence of tumor recurrence, whereas inflammation and immunosuppression did not. Here, we analyzed the cytotoxic CD8 T cells markers in relation with helper T subpopulation (Th1, Th2, Th17 and Treg).

Unsupervised hierarchical clustering of a correlation matrix revealed functional clusters of genes associated with Th17, Treg, Th2, Th1 and cytotoxic. Patients with high expression of the Th17 cluster had a poor prognosis whereas patients with high expression of the Th1 cluster had a prolonged disease-free survival. In contrast, no prediction of the prognosis was associated with Th2 or Treg clusters. The combined analysis of cytotoxic and Th17 clusters gave a better discrimination for relapse. These results were confirmed by in situ analysis.

## Results

We first investigated the expression of immune gene in colorectal tumors and adjacent normal mucosa from the same patient by real-time PCR. The gene expression levels of regulatory T cells (FoxP3, CTLA4, TGF $\beta$ 1) and cytotoxic T cells (IFNG, GZMB) were significantly higher in tumoral than in normal tissue (fig. 1). We observed the same tendency for T helper 2 cells (IL-4, IL-5, IL-13) and T helper 17 cells (IL-17A). On the contrary, the gene expression level of Th2 (IL4R, STAT6), Th1 (TBX21, IL12RB2 IRF1) and Th17 (RORC) were the same in tumoral and normal tissue.

Evaluation of intratumoral immune coordination was assessed by analyzing the correlation between 45 markers for immune cells populations (Th1, Th2, Treg, Th17 and cytotoxic T cells). Pairwise comparisons of gene expression level were done by measuring Pearson correlation coefficients ( $r$ ). The relationships implied by these correlations were visualized by using unsupervised hierarchical clustering of  $r$  values (fig. 2A). We identified 8 clusters of co-modulated genes. Two clusters were composed essentially with Th1 genes: Th1 cytotoxic cluster (IRF1, GZMB, IL27, GNLY, PRF1, CCL5, CD8A, STAT1) and Th1 (cluster A.2) (IL12RB1, CD28, CCR5, HLA DMB, IL12RB2, CD38, CXCR6, TBX21). We identified 3 clusters with essentially Th2 genes: Th2 cluster B.1 (IFNGR1, STAT3, IFNGR2, IL10RB, IL4R, STAT6), Th2 cluster B.2 (IL4, IL5, IL13) and Th2 cluster B.3 (CCR7, CD3E, CD40LG, CCL19, CCR4, GATA3). Two regulatory genes (IL10, TGF $\beta$ ) clustered together and a Treg cluster including FoxP3 and CTLA4, and the two chemokines (CCL17, CCL22) involved in Treg and Th2 process, was found. The last cluster was composed with Th17 genes (RORC, IL17A) and another chemokine (CCL24). In order to validate the gene expression coordination, we performed in situ double staining with anti CCL24 antibody and with anti RORC or IL17A antibodies. By immunofluorescence, cells RORC+ or IL17A+ were found to be CCL24+ (fig. Sup. 1). We performed disease-free survival analysis for up to 10 years after primary tumor resection (fig. 2B). We determined the optimal and the median cut-off values for each genes. Logrank P-value and hazard ratios associated with disease-free survival were then calculated. Most of genes from the Th1 cytotoxic (cluster A.1) and the Th1 (cluster A.2) had hazard ratio lower than 1 with the optimal and the median cut-off illustrating the beneficial effect of Th1 cytotoxic gene on survival. In contrast, the Th17 cluster had a hazard ratio higher than 1.

To better characterize the mechanisms involved in antitumoral activity, we investigated the effect on clinical outcome of mRNA expression level. For each cluster defined by the correlation matrix, patients were classified according to the expression levels of mRNAs (fig. 3 and fig. sup. 2). Groups of patients were defined using unsupervised hierarchical tree structure. High expression level of Th1 cytotoxic cluster genes was associated with significantly improved disease-free survival rates ( $p=0.01$ ) (Fig. 3A). Patients with low expression of this cluster had a early recurrence: Median of relapse were 18 months and 78 months for patients with low and high expression of Th1 cytotoxic cluster, respectively. In the same way, patient with high expression level of Th1 cluster A.2 seemed to have prolonged disease-free survival. The expression level of other gene-clusters (Th2, regulatory T clusters) had no apparent effect on clinical outcome (fig. 3B,C). On the contrary, patients with low expression level of Th17 cluster seemed to prolonged disease-free survival (fig. 3D). Thus, 80% of the patients with low expression did not experience relapse.

Patients were classified in 3 groups (lo, het and hi) according to the Th17 hierarchical tree structure. Gene expression of IL17A and RORC were used as controls (fig. 4 and fig. sup. 2). In lo and het Th17 cluster groups, the expression level of IL17A was very low and high in the Th17 cluster hi group. The expression level of RORC was low in the group Th17 cluster lo and high in the other groups. Gene expression of CCL24, IL23R, FOXP3, CCL22, IFNGR2 and STAT6 significantly increased with the genes expression of Th17 cluster, whereas gene expression of Th1 and cytotoxic gene (PRF1, TBX21, CD8A) significantly decreased.

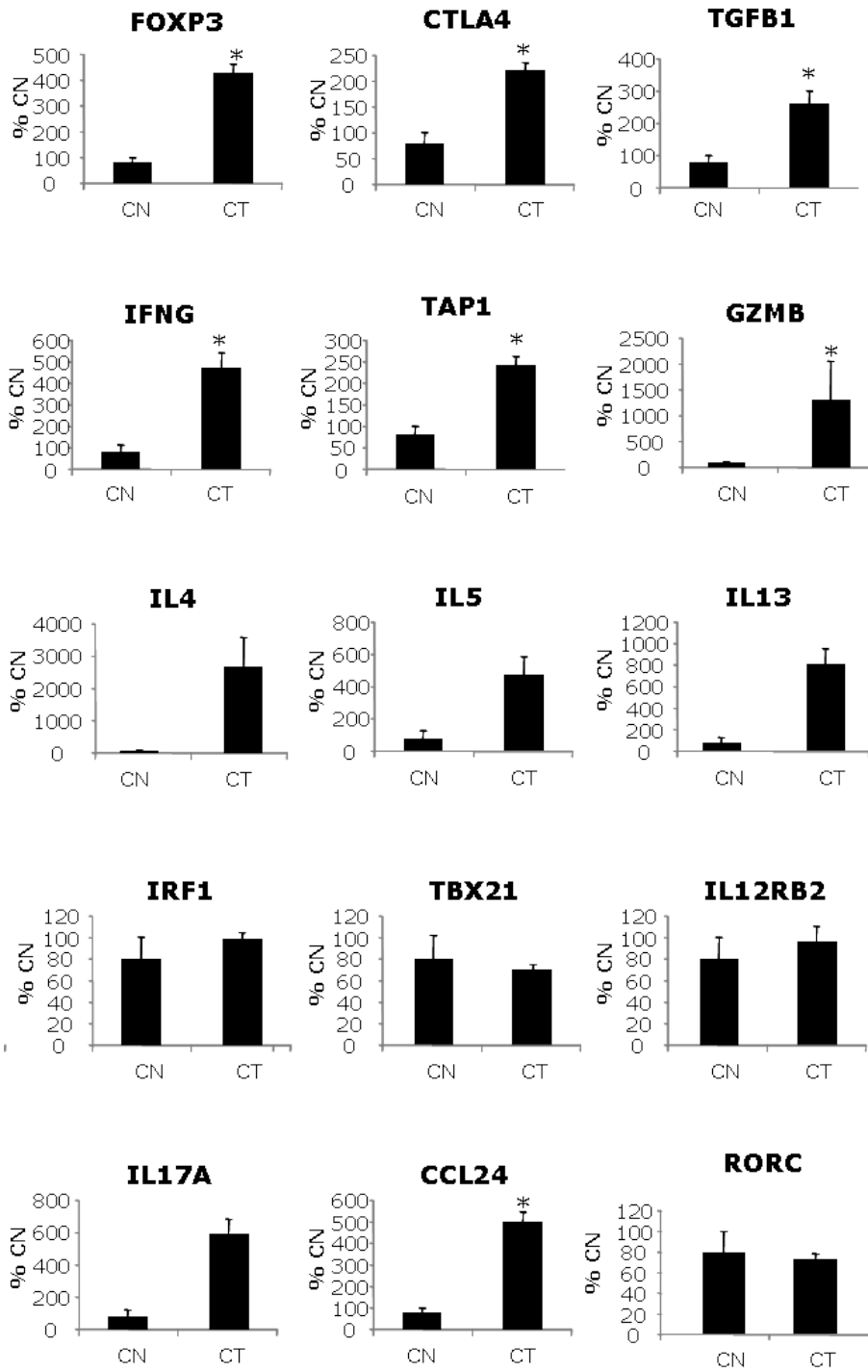


Figure 1: Gene expression level within primary colorectal tumors (n=126) compared to normal tissue (n=3). \*,  $p \leq 0.05$  wilcoxon test.

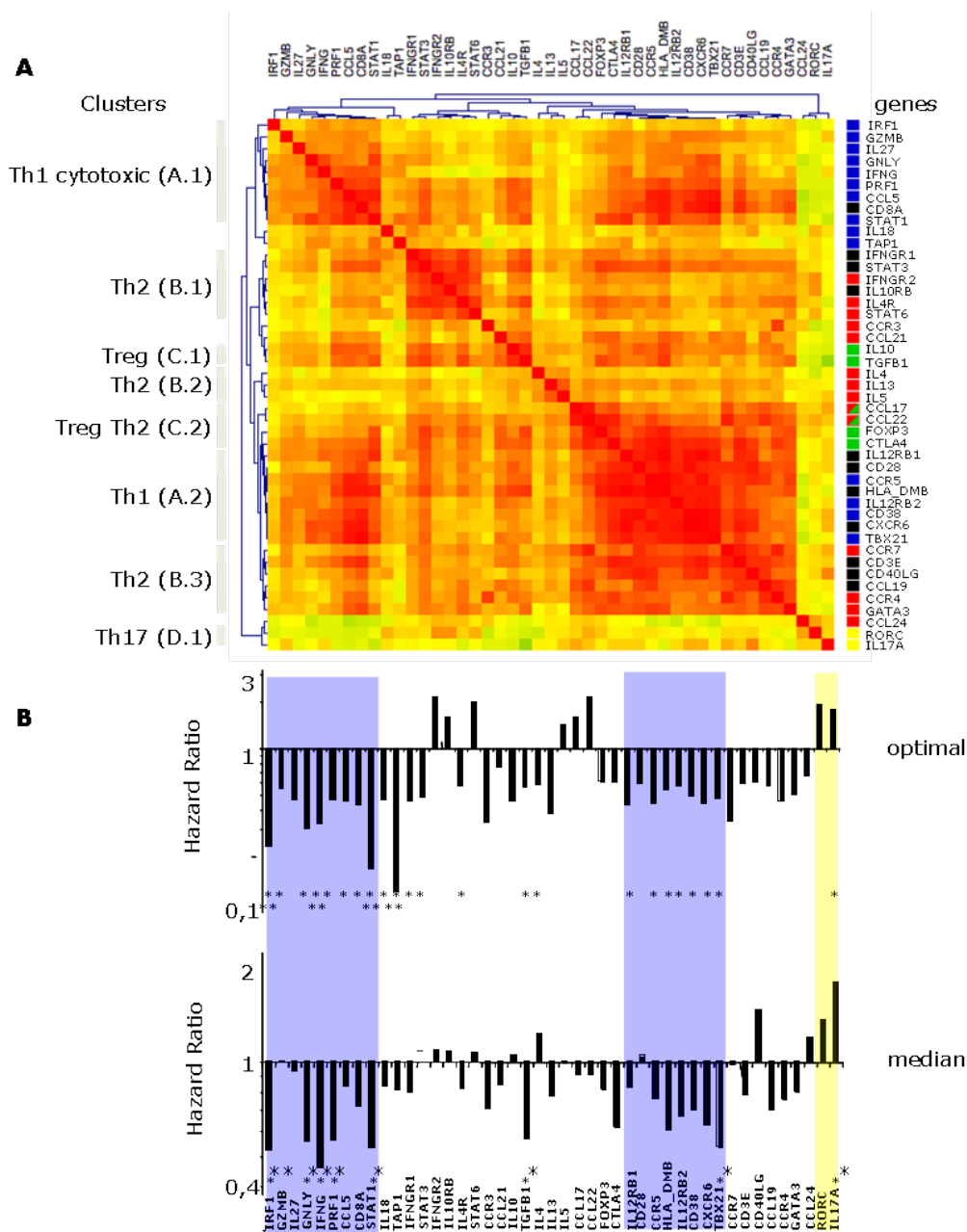


Figure 2: (A) Unsupervised hierarchical clustering of correlation matrix of mRNA expression level from 126 frozen tumors. Pearson correlation coefficients (R) were calculated between the combinations of 45 markers for immune cells populations. Correlation coefficients were plotted with negative correlation (green,  $R=-0.8$ ), positive correlation (red,  $R=0.8$ ) and  $R=0$  (yellow), in matrix representation followed by unsupervised hierarchical clustering. Th1 cells genes are represented in blue, Th2 cells genes in red, regulatory T cells genes in green, Th17 cells genes in yellow and total T cells in black. (B) Hazard ratio (HR) for DFS comparing patients with low against high gene expression. Genes were classified according to the correlation matrix. The cut point value for the expression of each gene was defined at the optimal (top) or the median (bottom) cutoff of the cohort. \* $p \leq 0.05$ , \*\* $p$ -corrected  $\leq 0.05$ .

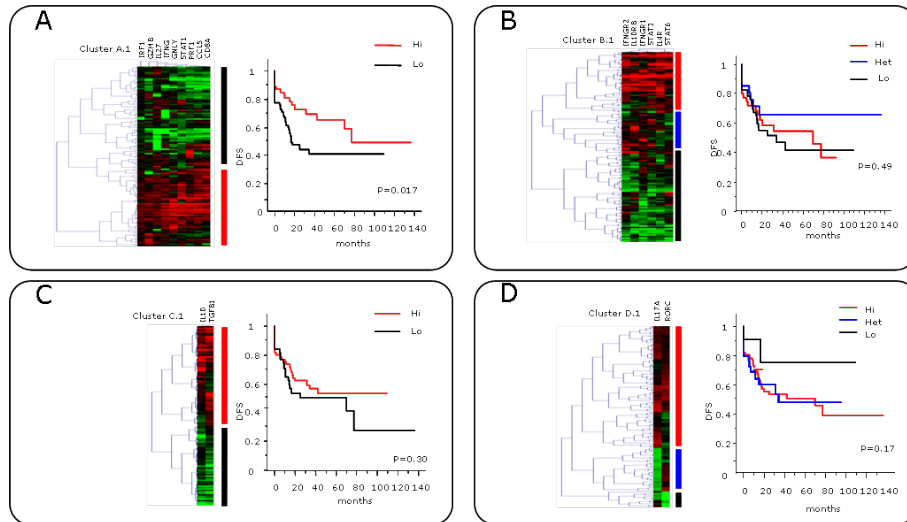


Figure 3: Hierarchical tree structure classifying the 126 patients according to the mRNA level of genes from the Th1 (A) Th2 (B) Treg (C) Th17 (D) clusters from maximal (red) to minimal (green) expression levels. Kaplan-Meier curves illustrated the duration of disease free survival according to the high (hi, red line) heterogeneous (het, blue) or low (lo, black line) gene expression.

We then evaluated the intratumoral expression of IL17A according to tumor progression within the colon mucosa (T-stage) and according to relapse status of the patients. The expression level of IL17A decreased with the T-stage (fig. 5A). Patients presenting an expression level of IL17A superior to the median of the cohort were named "hi" and otherwise "Lo". Six groups of patients was determined with the T-stage (Tis-2, T3, T4) and relapse status. The frequency of patients "Hi" (or "Lo") was calculated in each one of these 6 groups. In the early stages, the frequency of patients "Hi" was higher than in the late stages (fig. 5B). The frequency of patients "Hi" decreased with the stage, whereas the frequency of patients "Lo" increased. In a given T-stage, the frequency of patients "Hi" was higher in groups of patients with relapse than without relapse.

We assessed the complementary effect of Th1 and Th17 reaction. Kaplan-Meier curves illustrated the combination of groups from Th1 cytotoxic cluster (Hi and Lo) and the groups from Th17 cluster (Lo and Hi) (Fig. 6). In the group Th1 cytotoxic hi, patients with low expression of Th17 cluster genes had a better survival than patients with high expression level of Th17 cluster genes. Similarly in the group Th1 cytotoxic lo, patients with low expression of Th17 cluster genes had a better survival than patients with high expression level of Th17 cluster genes. Thus the best combination associated with prolonged disease-free survival was Th1-hi and Th17-low whereas the worst combination was Th1-lo and Th17-hi.

We validated our mRNA results with in situ studies with tissue microarray in the center and the invasive margin of the tumor. Immunostaining for Th1 (CD8, T-bet), Th17 (IL17) and regulatory T cells (FOXP3) were illustrated (fig. 7) and quantified with the use of dedicated image analysis workstation.

The percentage of patient with relapse decreased with the CD8 density (fig. 8A): 0%, 42% and 70% in CD8 HiHi, Het, LoLo patients groups, respectively. Contrary to CD8, the frequency of patient with relapse increased with the IL17 density (83%, 33%, 0% in IL17 HiHi, Het, LoLo and CD8 Het patients groups and 100%, 68%, 33% in IL17 HiHi, Het, LoLo and CD8 LoLo patients groups respectively). The density of FoxP3+ decrease with the CD8 density (fig. sup. 3). We evaluate disease-free survival according to the densities of Th1 and Th17 cells. Kaplan-Meier curves illustrated the pejorative effect

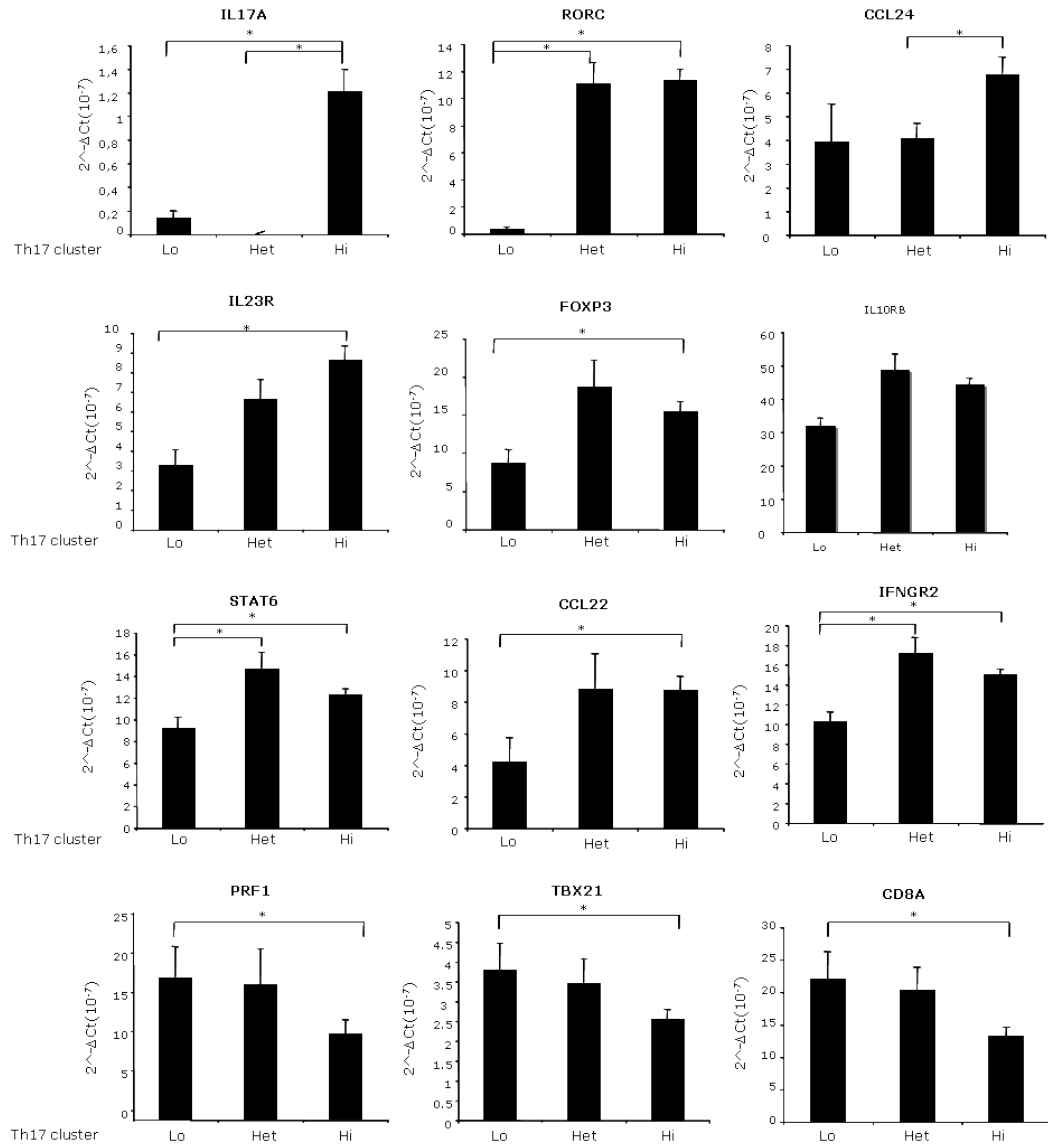


Figure 4: Gene expression level within primary colorectal tumors. Patients (n=126) were classified in three groups according to the Th17 hierarchical tree structure. Relative expression levels were adjusted for the level of 18S mRNA for each sample. The levels were represented as mean of 2-dCt values. \*,  $p < 0.05$  wilcoxon test

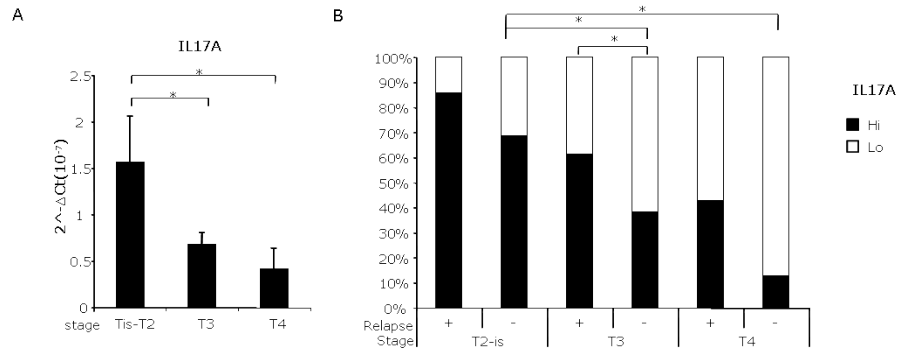


Figure 5: Gene expression level of IL17A within colorectal tumors. Patients (n=126) were classified according to the T-stage (**A**). (**B**) frequency of patients with high (black columns) and low (white columns) expression of IL17A classified according to T-stage and relapse. The cutoff value for the expression of IL17A was defined at the median of the cohort. \*,  $p \leq 0.05$  wilcoxon test

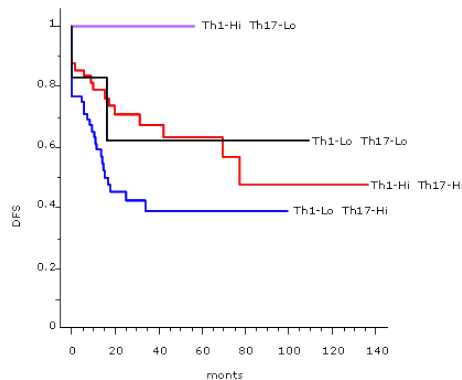


Figure 6: Disease free survival of colorectal cancer patients according to the expression level of the Th1 cytotoxic gene cluster in combination with Th17 genes cluster (red Th1-Hi Th17-Hi, black Th1-Lo Th17-Lo, blue Th1-Lo Th17-Hi, purple Th1-Hi Th17-Lo)



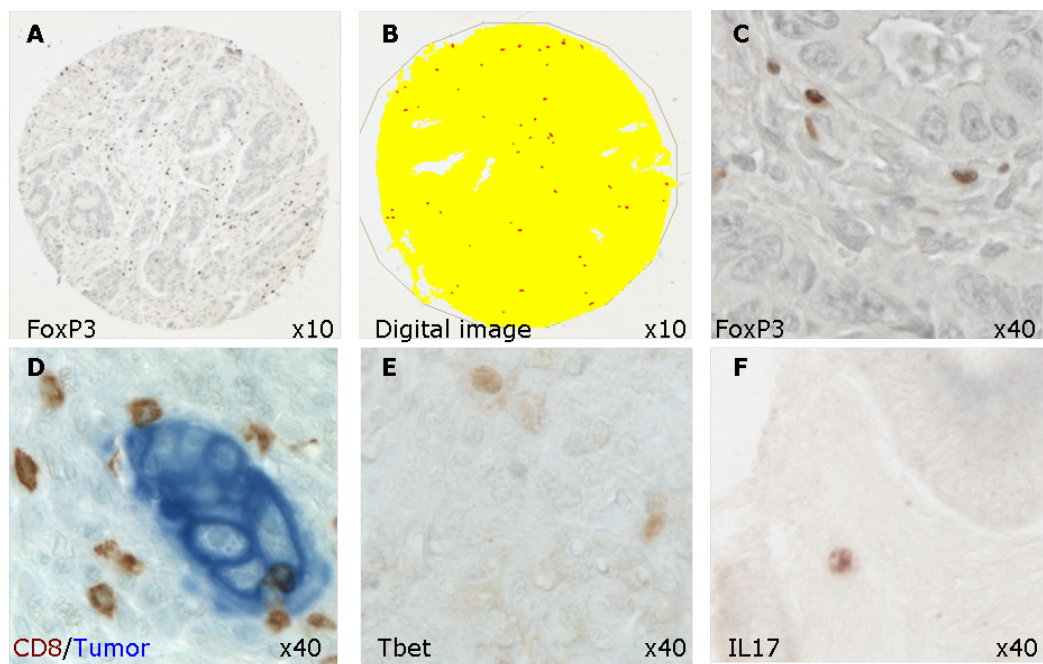


Figure 7: A representative example of FoxP3 (**A**, **C**), CD8 (**D**), Tbet (**E**) and IL17 (**F**) immunostaining of colorectal cancer tissue microarray. Digital image analyzed with the image software SpotBrowser with tissue represented in yellow and the FoxP3 cells represented in red (middle). The densities of this immune cells were recorded as the number of positive cells per unit of tissue surface area ( $\text{mm}^2$ ).

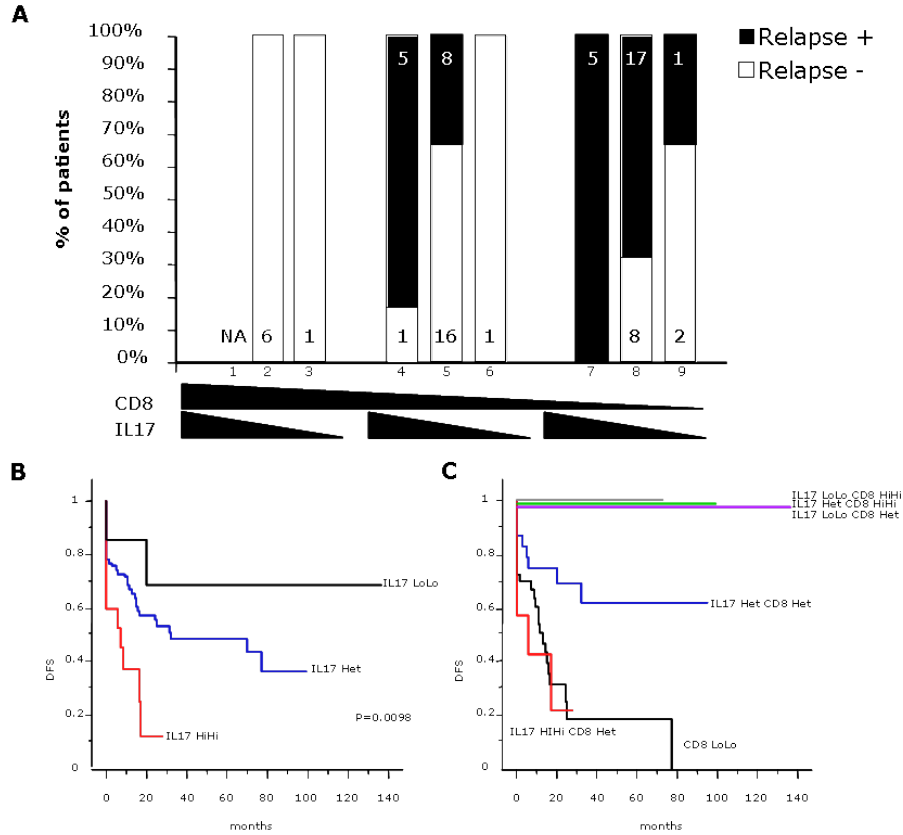


Figure 8: **(A)** Frequency of patients with (black columns) or without (white columns) relapse classified according to the densities of immune cells within colorectal tumors. Patients (n=71) were classified according to the CD8 and IL17 cells densities in two region of the tumor, the center and the invasive margin of the tumor. High densities in the two regions of CD8 were represented in the histograms 1-3, heterogeneous densities in the histograms 4-6, and low densities in the histograms 7-9. High densities in the two regions of IL17 were represented in the histograms 1,4,7, heterogeneous densities in the histograms 2,5,8, and low densities in the histograms 3,6,9.

Kaplan Meier curves illustrate the duration of Disease free survival according to the IL17 cells density **(B)** in combination with CD8 cells density **(C)** in two region of the tumor, the center and the invasive margin of the tumor. Groups of patients with low densities in the two regions were named LoLo, those with high densities in the two regions HiHi, and the others with heterogeneous densities Het. The cutoff value for the density of CD8 and IL17 cells were defined at the optimal p-value of the cohort.

of IL17A on patient's survival. Patients with low density of IL17 in the two regions of the tumors had a better disease-free survival than patients with high density of IL17 (fig. 8B). Indeed, 70% of the patients in the LoLo group did not experience relapse. In contrast, 50% of the patients in the HiHi group had tumor recurrence after one year. On the contrary, patients with high density of CD8, T-Bet or FOXP3 had a better disease-free survival than patients with low density (fig. sup. 4). The combination of the CD8 and IL17 markers defined groups of patients with very different outcome. Patients with CD8 LoLo had a dramatic outcome. Median disease-free survival was 9 months and all patients experienced tumor recurrence after 6 year. In contrast, patient with CD8 HiHi had a very good outcome whether they had heterogeneous (Het) or low (LoLo) densities of IL17. None of this patients experienced relapse. Interestingly, no patient presented with CD8 HiHi and IL17 HiHi. Strikingly, in the group of patients with heterogeneous density of CD8, patients with low density of IL17 had a better disease-free survival whereas patients with high density of IL17 had a poor prognosis (fig. 8C and fig. sup. 5). There was no relapse in the group of patient with low density of IL17 cells whereas 80% of patients in the high density group relapsed.

## Discussion

We analyzed T helper subpopulations in colorectal cancer. We showed that immune markers expression were not similar in normal and tumoral tissues. Unsupervised hierarchical clustering of a correlation matrix revealed functional cluster of immune genes: Th17, Th1, Th2, Treg and cytotoxic. Thus, immune coordination between T helper subsets was found within the tumors. Th1, cytotoxic and Th17 clusters were associated with a particular clinical outcome: Patients with high expression of the Th17 cluster genes had a poor prognosis whereas patients with high expression of Th1 or cytotoxic cluster genes had a prolonged disease-free survival. The combined analysis of the cytotoxic and Th17 clusters gave a better discrimination for relapse. These results were confirmed by in situ analysis: patients with high density of IL17 cells had a poor prognosis whereas patients with high density of CD8 cells had a prolonged disease-free survival. The combination of these two markers gave also a better discrimination of patients. We previously showed that Th1 pattern was associated with increased expression of CX3CL1 ([6]) and memory T cell response with increased CXCL9 and CXCL10 expression ([6]). Here we show for the first time that Th17 produced CCL24, a chemokine known to affect eosinophil. We have shown that High Th1 and cytotoxic genes level were associated with a good outcome. These results have been validated with the CD8 and T-bet protein studies that confirms previous studies [4] [5] [7] [8] [9]. In contrast to Th1, we have shown that patients with a low Th17 response had a better disease-free survival. However the Th17 response decrease with the tumor progression in the tissue (T-stage). On the contrary, the Th17 response increase with advanced gastric cancer [10]. Kryczek et al [11] observe the opposite effect in ovarian cancer: tumor infiltrating Th17 are reduced in more advanced diseases and low level of IL17 is associated with a bad outcome . They also show that Th17 are positively associated with effectors cells. In our study, we did not found patient with high densities of IL17 and CD8 positive cells in situ. Patients with high density of CD8 did not experience relapse whatever the density of IL17. Thus, the positive effect of Th1 seem to be superior to the negative effect of Th17 on patient's survival. In contrast, in patients with heterogeneous densities of CD8 T cells between tumor regions, IL17 had a strong impact on patient's survival. We found three different clusters associated with Th2 response, but none of this cluster was associated with a particular outcome. Previously, Th2 response is associated with the tumor immune evasion [12] [13], or not associated [14] [15]. We found two different clusters with regulatory genes, including cytokines, chemokines and transcription factors that were not associated with a particular outcome. At the protein level, patients with high density of FoxP3 positive cells had a better disease-free survival. This result confirms the analysis of the FoxP3 mRNA expression: high level seems to be better for patient's survival. The role of regulatory T cells in cancer is disputed. In breast cancer, FoxP3 expression is associated with the worst overall survival [16] [17]. In colorectal cancer, Loddenkemper et al [18] don't find any difference between patients with high or low Treg infiltration whereas Salama et al [19] found an improved survival associated with a high density of tumor infiltrating FoxP3.

## Conclusion

In summary, unsupervised hierarchical clustering of a correlation matrix revealed functional clusters of genes associated with Th17, Treg, Th2, Th1 and cytotoxic. Patients with high expression of the Th17 cluster had a poor prognosis whereas patients with high expression of the Th1 cluster had a prolonged disease-free survival. The combined analysis of cytotoxic and Th17 clusters gave a better discrimination for relapse. These results were confirmed by in situ analysis.

## materials and methods

### Patients and database

Patients with colorectal cancer who underwent a primary resection at the Laennec/HEGP Hospital between 1986 and 2004 were randomly selected. Time to recurrence or disease-free time was defined as the time period from the date of surgery to confirmed tumor relapse date for relapsed patients and from the date of surgery to the date of last follow-up for disease-free patients. A secure Web-based database, TME.db, integrates the clinical data and the data from high-throughput technologies.

### Real-time PCR analysis

Tissue samples were snap-frozen. Total RNA was extracted by homogenization with RNeasy isolation kit (Qiagen). The integrity and the quality of the RNA (n=126) were evaluated on Bioanalyzer-2100 (Agilent Technologies). Quantitative real-time PCR was performed using Low-Density-Arrays and the 7900 robotic real-time PCR-system (Applied-Biosystems) according to the manufacturer's instructions. 18S primers and probes were used as internal controls. Data were analyzed using the SDS Software v2.3 (Applied-Biosystems).

### Construction of tissue microarrays

Using a tissue microarray instrument (Beecher Instruments, Alphelys, Plaisir, France), we removed two different and representative areas of the tumor: center of the tumor (CT) and invasive margin (IM) were punched (0.6 mm and 1 mm-diameter, respectively) from paraffin-embedded tissue-blocks. Tissue microarrays were cut into 5- m sections for immunohistochemical staining.

### Immunohistochemistry

After antigen retrieval and quenching of endogenous peroxidase activity, sections were incubated for 60 min at room temperature with antibodies against CD8 (4B11; Neomarkers, Fremont, CA), FoxP3 (ab20034; abcam, Cambridge, United Kingdom), T-bet (4B10) and IL17 (H-132; Santa Cruz Biotechnology, Santa Cruz, CA). Envision+ system (enzyme-conjugated polymer backbone coupled to secondary antibodies) and DAB-chromogen were applied (Dako, Copenhagen, Denmark). Tissue sections were counterstained with Harris's hematoxylin. Slides were analyzed using an image analysis workstation (SpotBrowser, Alphelys, Plaisir, France). The density was recorded as the number of positive cells per tissue surface area. For each duplicate, the mean density was used for statistical analysis.

### Statistical analysis

Genesis clustering software was used to visualize the correlation matrix, genes expression level matrix and to perform covariance value or Pearson un-centered hierarchical clustering. Kaplan-Meier curves were used to assess the influence of immune parameters on disease-free survival. The significance of these parameters was assessed by univariate analysis with the use of log-rank test. For pairwise comparisons Wilcoxon rank-sum test was used.  $P < 0.05$  was considered statistically significant. All analyzes were performed with the statistical software R and Statview.

# Bibliography

- [1] L. M. Coussens, Z. Werb. Inflammation and cancer. *nature*, 420:860–867, 2002.
- [2] F. Balkwill, L.M. Coussens. Cancer: an inflammatory link. *nature*, 431:405–407, 2004.
- [3] F. Pagès, A. Berger, M. Camus, F. Sanchez-Cabo, A. Costes, R. Molidor, B. Mlecnik, A. Kirilovsky, M. Nils-son, D. Damotte, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer, 2005.
- [4] J. Galon, A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pages, M. Tosolini, M. Camus, A. Berger, P. Wind, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome, 2006.
- [5] M. Camus, M. Tosolini, B. Mlecnik, F. Pages, A. Kirilovsky, A. Berger, A. Costes, G. Bindea, P. Charoentong, P. Bruneval, et al. Coordination of Intratumoral Immune Reaction and Human Colorectal Cancer Recurrence. *Cancer research*, 69(6):2685, 2009.
- [6] B. Mlecnik, M. Tosolini, P. Charoentong, A. Kirilovsky, G. Bindea, A. Berger, M. Camus, M. Gillard, P. Bruneval, W.H. Fridman, F. Pagès, Z. Trajanoski, J. Galon. Biomolecular network reconstruction identifies T cell homing factors associated with survival in colorectal cancer. *Gastroenterology*, page in press, 2009.
- [7] K. M. Ropponen, M.J. Eskelinen, et al. Prognostic value of tumour-infiltrating lymphocytes (TILs) in colorectal cancer. *J. Pathol.*, 182:318–324, 1997.
- [8] M. Guidoboni, R. Gafa, et al. Microsatellite instability and high content of activated cytotoxic lymphocytes identify colon cancer patients with a favorable prognosis. *American journal of pathology*, 159:297–304, 2001.
- [9] HE Lee, SW Chae, YJ Lee, MA Kim, HS Lee, BL Lee, WH Kim. Prognostic implications of type and density of tumour-infiltrating lymphocytes in gastric cancer. *British journal of cancer*, 99(10):1704–1711, 2008.
- [10] B. Zhang, G. Rong, et al. The prevalence of Th17 cells in patients with gastric cancer. *Biochem Biophys Res Commun*, 374:533–537, 2008.
- [11] I. Kryczek, M. Banerjee, et al. Phenotype, distribution, generation, and functional and clinical relevance of Th17 cells in the human tumor environments. *Blood*, 114(6):1141–1149, 2009.
- [12] E. Osawa, Nakajima, et al. Predominant T helper type 2-inflammatory responses promote murine colon cancers. *International Journal of Cancer*, 118(9):2232–2236, 2006.
- [13] A. Ziegler, R. Heidenreich, et al. EpCAM, a human tumor-associated antigen promotes Th2 development and tumor immune evasion. *Blood*, 113(15):3494–3502, 2009.
- [14] R.A. Kemp, F. Ronchese. Tumor-Specific Tc1, But Not Tc2, Cells Deliver Protective Antitumor Immunity 1. *The Journal of Immunology*, 167(11):6497–6502, 2001.
- [15] C. F. Evans, C. Galustian, et al. The effect of colorectal cancer upon host peripheral immune cell function. *colorectal disease*, 2009.
- [16] A. Merlo, P. Casalini, et al. FOXP3 expression and overall survival in breast cancer. *Journal of clinical oncology*, 27(11):1746–1752, 2009.
- [17] S.F. Adams, D.A. Levine, et al. Intraepithelial T cells and tumor proliferation: impact on the benefit from surgical cytoreduction in advanced serous ovarian cancer. *Cancer*, 115(13):2891–2902, 2008.

- [18] C. Loddenkemper, M. Schernus, et al. In situ analysis of FOXP3+ regulatory T cells in human colorectal cancer. *Journal of translational medicine*, 4(52), 2006.
- [19] P. Salama, M. Phillips, et al. Tumor-infiltrating FOXP3+ T regulatory cells show strong prognostic significance in colorectal cancer. *Journal of clinical oncology*, 27(2):186–192, 2009.

## Supplemental figure



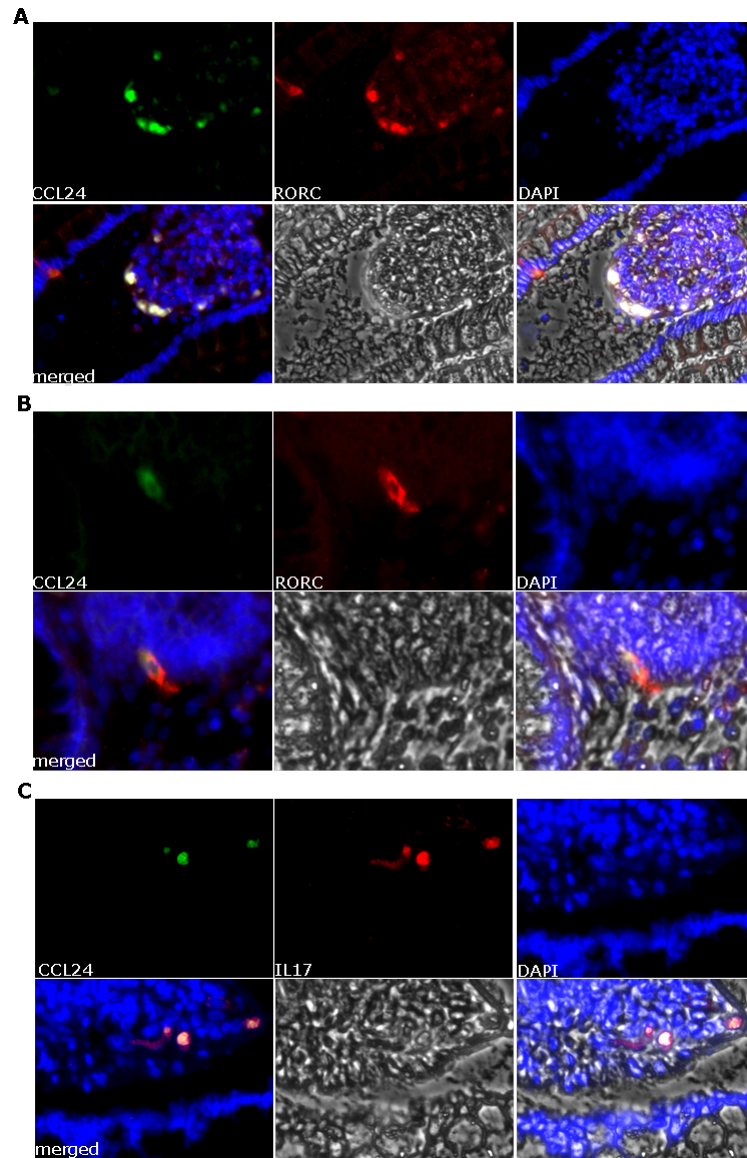


Figure 1: A representative example of CCL24/RORC (A, B), and CCL24/IL17 (C) immunostaining of colorectal cancer

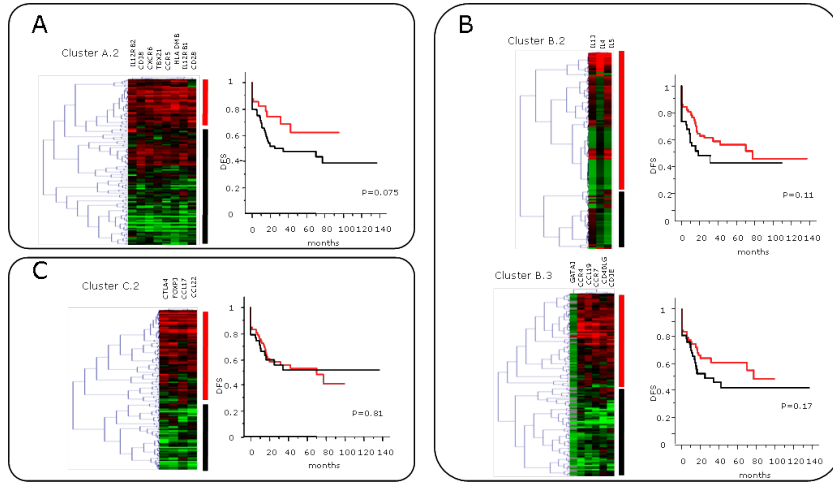


Figure 2: Hierarchical tree structure classifying the 126 patients according to the mRNA level of genes from the Th1 (A) Th2 (B) Treg (C) clusters from maximal (red) to minimal (green) expression levels. Kaplan-Meier curves illustrated the duration of disease free survival according to the high (Hi, red line) or low (lo, black line) gene expression.

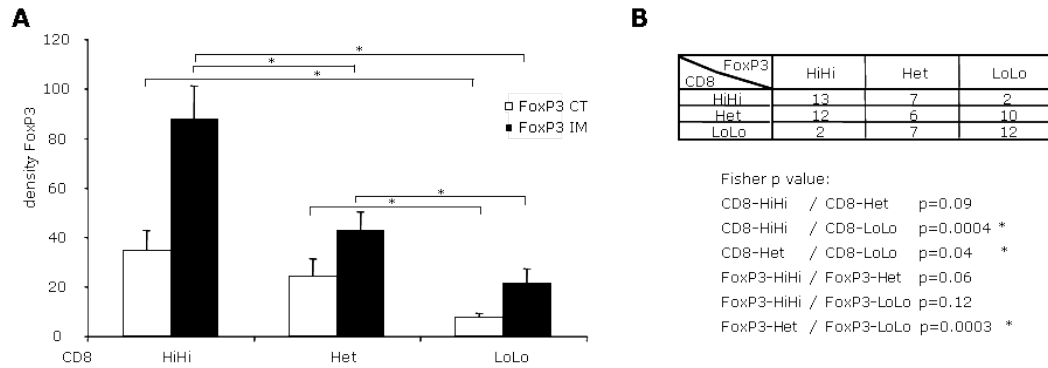


Figure 3: (A) Density of FoxP3 in two region of the tumor, the center (white columns) and the invasive margin (black columns). Patients (n=71) were classified according to the CD8 cells densities in the two region of the tumor. \*,  $p \leq 0.05$  wilcoxon test (B) Patients (n=71) were classified according to the CD8 and FoxP3 densities in the two regions of the tumor. \*,  $p \leq 0.05$  Fisher test The cutoff value for the expression of IL17A was defined at the median of the cohort.

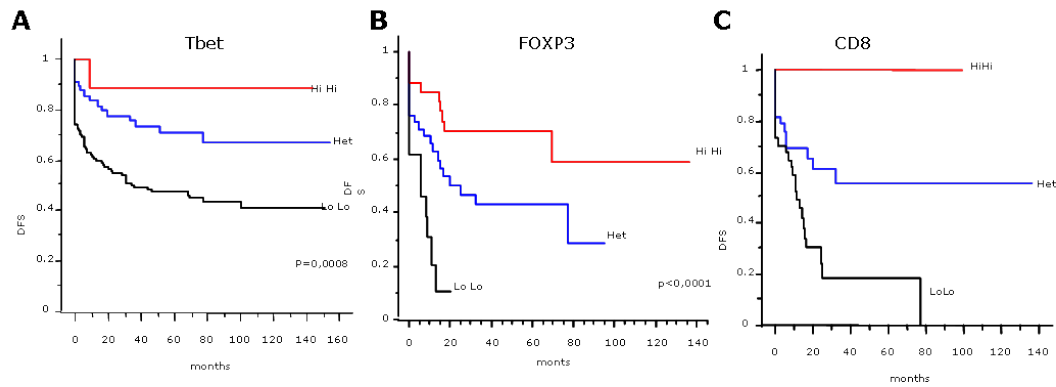


Figure 4: Kaplan Meier curves illustrate the duration of Disease free survival according to the T-Bet (A), FoxP3 (B) and CD8 (C) cells density in two region of the tumor, the center and the invasive margin of the tumor. The cutoff values were defined at the optimal p-value of the cohort.

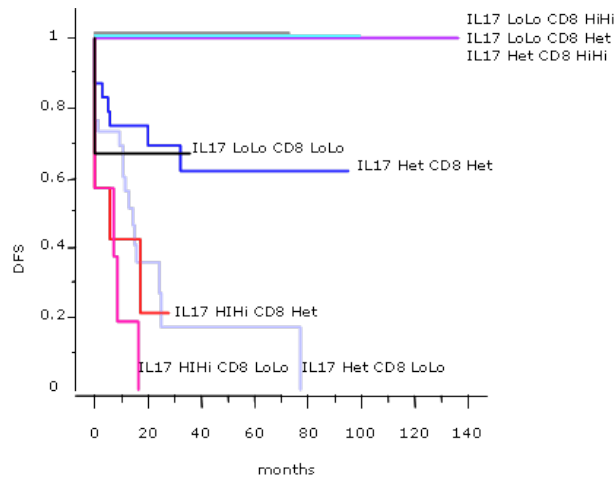


Figure 5: Kaplan Meier curves illustrate the duration of Disease free survival according to the IL17 cells density in combination with CD8 cells density in two region of the tumor, the center and the invasive margin of the tumor. The cutoff value for the density of IL17 and CD8 cells was defined at the optimal p-value of the cohort.

# Données non publiées

## Les lymphocytes T régulateurs

Les lymphocytes T régulateurs contrôlent la réponse immunitaire. Je me suis donc intéressée à cette population au site tumoral en les étudiant par immunohistochimie.

## Matériel et méthodes

### Construction des Tissue MicroArrays

Des Tissus MicroArrays (TMA) ont été construits en ciblant deux zones représentatives de la tumeur : le centre de la tumeur (CT) et le front de progression (FP). Les carottes de 0,6 et 1 mm de diamètre pour, respectivement le centre et le front de progression, ont été faites dans des blocs de tissu en paraffine et insérées dans un nouveau bloc de paraffine. Les TMA ont été ensuite coupés en section de 5 $\mu$ m pour les marquages immunohistochimiques.

### Immunohistochimie

Après le démasquage antigénique à pH8, l'inhibition des peroxidases endogènes et la saturation, les coupes ont été incubées une heure à température ambiante avec les anticorps CD45RO (OPD4), mast cell tryptase (AA1), granulocyte (BM-2), CD3 (SP7 ; neomarker, Fremont, CA), FoxP3 (ab20034 ; abcam, Cambridge, Angleterre), CD68 (PGM1), CD20 (L26 ; Dako Copenhagen, Danemark), IL3RA (IL3RA), IL23 (IL23 ; Atlas Antibodies, Stockholm, Suède) et IL17 (H-132 ; Santa Cruz Biotechnology, Santa Cruz, CA). Le système Envision+ (polymères couplés à une enzyme) et le chromogène DAB sont ensuite appliqués (Dako, Copenhagen, Danemark). Les lames sont analysées à l'aide d'une station d'analyse d'images (SpotBrowser, Alphelys, Plaisir, France). La densité des cellules positives est calculée en effectuant le rapport nombre de cellules positives sur surface de tissu. Pour chaque patient, la moyenne des différents spots dans une région de la tumeur est calculée et utilisée pour les analyses ultérieures.

### Analyses statistiques

Le logiciel Genesis a été utilisé pour visualiser les matrices de corrélation et pour réaliser le classement hiérarchique "Pearson un-centered". Les courbes de Kaplan-Meier ont été utilisées pour

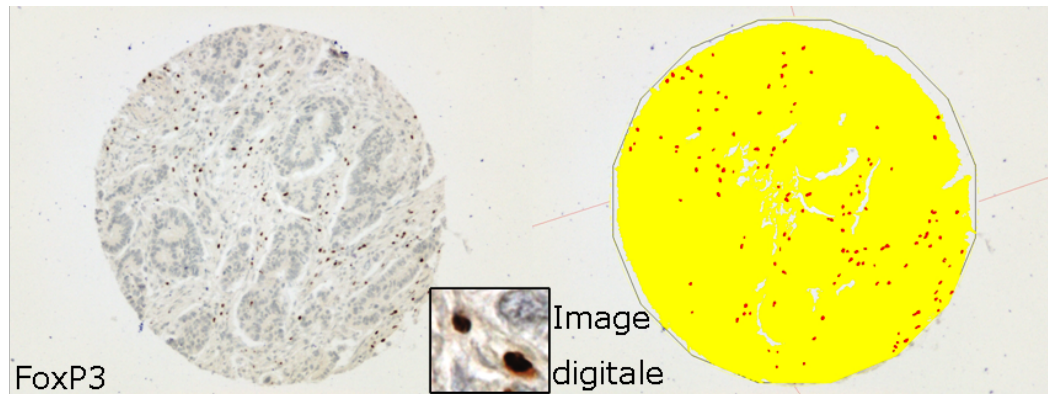


FIGURE 11 – Un exemple de marquage de cellules FoxP3+ en marron (gauche) et son analyse digitale (droite) effectué par le logiciel d’analyse d’image SpotBrowser. Les cellules FoxP3+ sont représentées en rouge et le tissu en jaune. La densité de cellules FoxP3+ est calculé en effectuant le rapport nombre de cellules positives sur surface de tissu (mm<sup>2</sup>).

visualiser l’influence des paramètres sur la survie sans récurrence des patients.  $p \leq 0,05$  a été considéré comme statistiquement significatif. Toutes ces analyses ont été faites avec les logiciels statistiques R et Statview.

## Résultats

### Les lymphocytes T régulateurs

J’ai effectué le marquage des cellules FoxP3+ et analysé les images obtenues (Figure 11). Le logiciel SpotBrowser compte le nombre de cellules FoxP3 (représentées en rouge) ainsi que la surface du tissu (représenté en jaune). Ses données permettent de calculer la densité de cellules dans chaque spot de tissu. Une moyenne des différents spots de chaque région de la tumeur pour chaque patient est ensuite calculée. Des données ont été obtenues pour 464 patients au centre de la tumeur et pour 471 patients au front de progression. 428 patients possèdent des données pour les deux régions de la tumeur.

Des courbes de survie sans récurrence ont été tracées afin d’étudier l’effet de ces cellules sur la récurrence des patients. Pour cela, les patients sont classifiés en groupes fortement (Hi) ou faiblement infiltrés (Lo) suivant leurs densités de cellules FoxP3+, en fonction de la médiane du groupe (figure 12).

Le groupe de patients fortement infiltrés (courbe rouge) en cellules FoxP3+ dans le centre de la tumeur ont une meilleure survie sans récurrence que le groupe de patients faiblement infiltrés (courbe noire,  $p \leq 0,0001$ ). Cinq ans (60 mois) après la chirurgie, 80% des patients fortement infiltrés n’ont pas récidivé contre 55% dans le groupe des patients faiblement infiltrés. Au front de progression, on peut observer la même tendance même si la différence entre les deux courbes est moins marquée et non significative ( $p=0,051$ ).

En combinant les groupes obtenus dans les deux zones de la tumeur, quatre groupes sont obtenus : deux groupes où les densités sont concordantes dans les deux régions et deux groupes où les densités

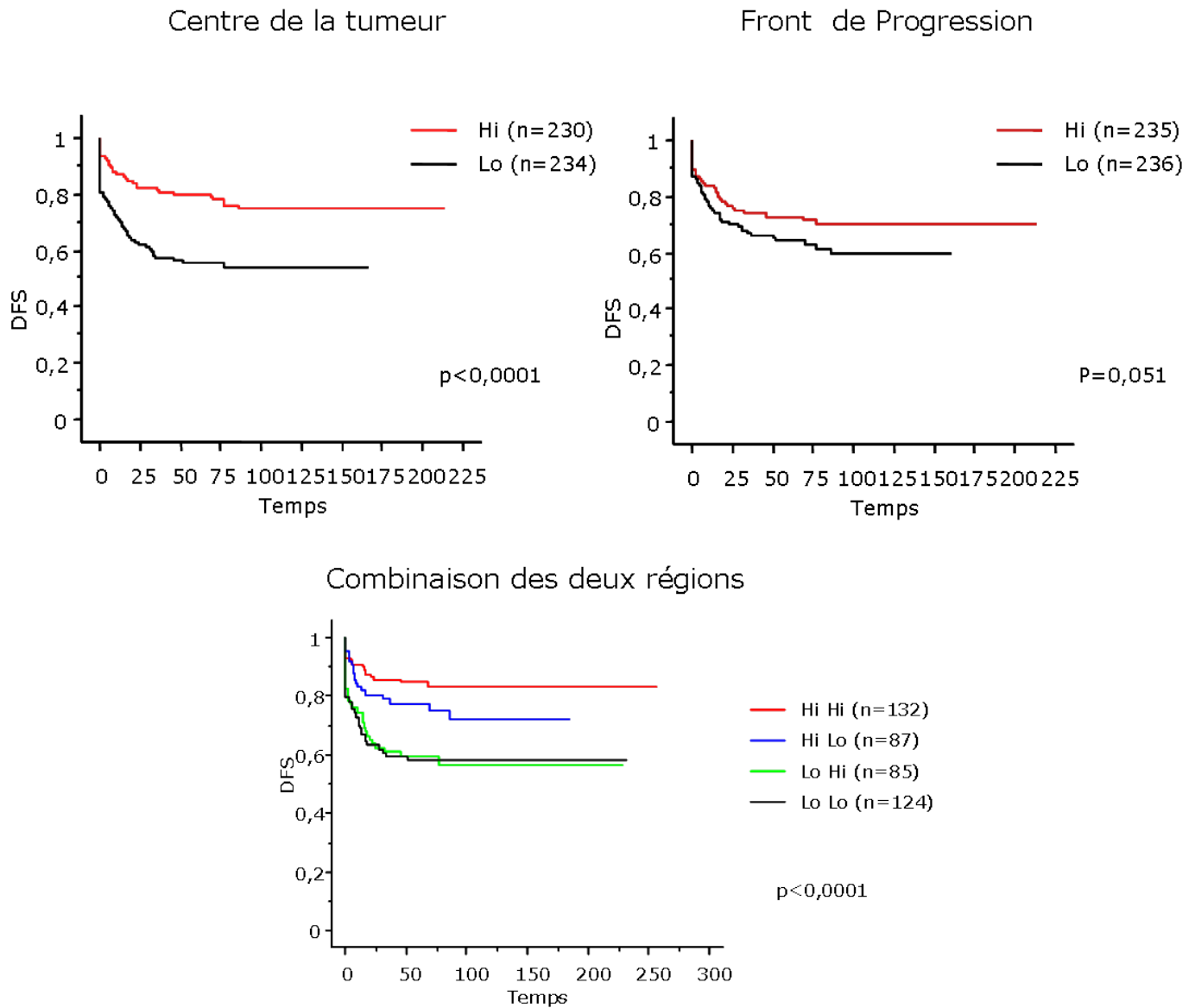


FIGURE 12 – Courbes de survie sans récurrence associées à la densité de cellules FoxP3+. Pour chaque région de la tumeur, les patients sont classés en groupe fortement (Hi, courbe en rouge) ou faiblement (Lo, courbe en noir) infiltrés par rapport à la médiane. Ces groupes sont ensuite combinés. Les patients fortement infiltrés (Hi Hi) dans les deux régions sont représentés en rouge, ceux faiblement infiltrés (Lo Lo) dans les deux régions en noir, ceux fortement infiltrés dans le centre de la tumeur et faiblement dans le front de progression (Hi Lo) en bleu, et ceux faiblement infiltrés dans le centre de la tumeur et fortement dans le front de progression (Lo Hi) en vert. La survie sans récurrence de ces groupes est visualisée par les courbes de Kaplan-Meier.

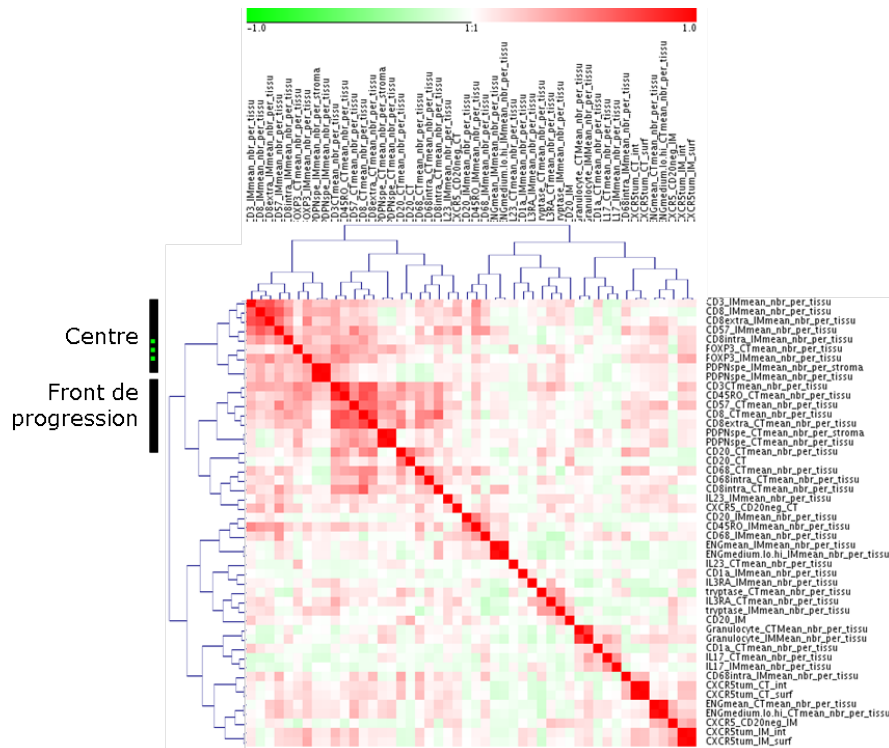


FIGURE 13 – Matrice de corrélation avec classement hiérarchique non supervisé des densités de cellules immunitaires au centre (CT) et front de progression (IM, Invasive Margin) de la tumeur. Les coefficients de corrélation de Pearson sont visualisés en vert pour une corrélation négative ( $R=-1$ ), en rouge pour une corrélation positive ( $R=1$ ) et en blanc pour une absence de corrélation ( $R=0$ ).

sont discordantes. Le groupe de patients fortement infiltrés dans les deux régions (courbe rouge) en cellules FoxP3+ ont une meilleure survie sans récurrence que le groupe de patients faiblement infiltrés dans les deux régions (courbe noire,  $p=0,005$ ). Le groupe de patients faiblement infiltrés dans le centre de la tumeur et fortement infiltrés dans le front de progression (Lo Hi, courbe verte) a une courbe de survie sans récurrence superposable à celle du groupe de patients faiblement infiltrés dans les deux régions. Le groupe de patients fortement infiltrés dans le centre de la tumeur et faiblement infiltrés dans le front de progression (Hi Lo, courbe bleue) ont une courbe de survie intermédiaire. Cinq ans (60 mois) après la chirurgie, le pourcentage des patients sans récurrence est respectivement de 85%, 75% et 60% pour les patients fortement infiltrés dans les deux régions, ceux fortement infiltrés dans le centre de la tumeur et faiblement dans le front de progression et pour ceux faiblement infiltrés dans le centre de la tumeur quelque soit l'infiltration au front de progression. Ces différences sont significatives ( $p=0,005$ ).

Une forte densité de cellules FoxP3+ est de bon pronostic, surtout au centre de la tumeur. Une forte densité au front de progression semble également nécessaire en complément du centre de la tumeur.

### Les lymphocytes T régulateurs et les lymphocytes T

Dans le but d'étudier la coordination des cellules FoxP3+ avec les autres types de cellules immunitaires étudiées, une matrice de corrélation a été réalisée (figure 13). Les corrélations ont été calculées

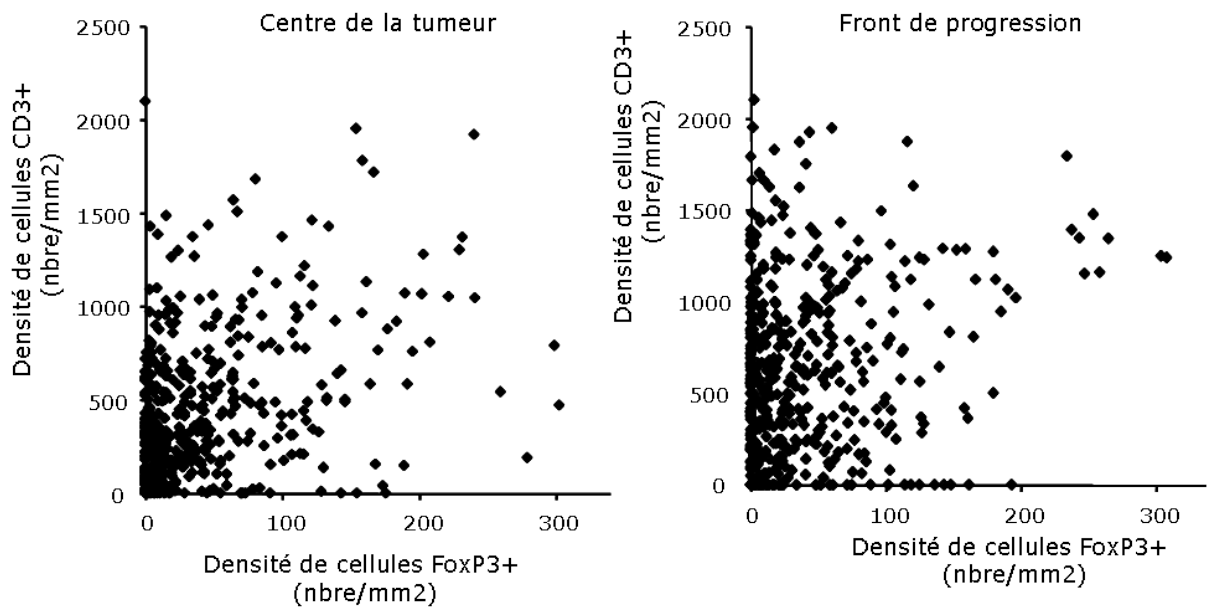


FIGURE 14 – Représentation graphique de la densité de cellules CD3+ et FoxP3+ dans chaque région de la tumeur. Chaque point représente un patient.

entre les différentes densités des cellules immunitaires dans le centre (CT) et le front de progression (IM) de la tumeur. La matrice est visualisée à l'aide du logiciel Genesis. Les fortes corrélations positives sont visualisées en rouge, les fortes corrélations négatives en vert et l'absence de corrélation en blanc. La matrice est classée hiérarchiquement de manière non supervisée.

On observe deux grands groupes de cellules au sein de la matrice. Dans chaque groupe, les cellules immunitaires proviennent de la même région de la tumeur. La première (en haut, à gauche) est constituée de densité provenant majoritairement du front de progression, à l'exception du marqueur FOXP3 dont les deux zones sont représentées. On y retrouve les marqueurs lymphocytaires CD3, CD8, CD57 et FoxP3 ainsi que les vaisseaux lymphatiques (podoplanine). Le deuxième cluster contient les mêmes marqueurs, au front de progression.

Les cellules FoxP3+ font partie d'un cluster comprenant des lymphocytes T au centre de la tumeur. Le bon pronostic associé à la densité de FoxP3 pourrait être la conséquence d'une corrélation avec la présence de lymphocytes T cytotoxiques ou Th1. Or, les graphiques ne montrent pas de relation nette entre la densité de cellules FoxP3+ et celle de cellules CD3+ (figure 14). Le rapport des densités de cellules FoxP3+ sur celles de cellules CD3+ le confirme (figure 15). Il est variable. En effet, certains patients présentent des cellules CD3+ sans cellule FoxP3+ (rapport égal à 0%), tandis que d'autres patients ont jusqu'à plus de six cellules FoxP3+ pour dix cellules CD3+ (rapport supérieur à 60%).

Les courbes de survie sans récurrence représentant les groupes définis à partir des ratios des densités des cellules FoxP3+ sur celles des cellules CD3+ confirme l'effet de la présence de lymphocytes T régulateurs (figure 15). En effet, les courbes définies par la densité de cellules FoxP3+ et définies par le ratio FoxP3/CD3 sont semblables. Dans le centre de la tumeur, les patients avec un ratio élevé ont une meilleure survie sans récurrence que les patients avec un ratio faible ( $p \leq 0,0001$ ). Le taux de survie sans récurrence à 5 ans est de 80% et 60% pour respectivement le groupe avec un fort et un faible ratio. La même tendance est observée au front de progression bien que ce ne soit pas significatif.



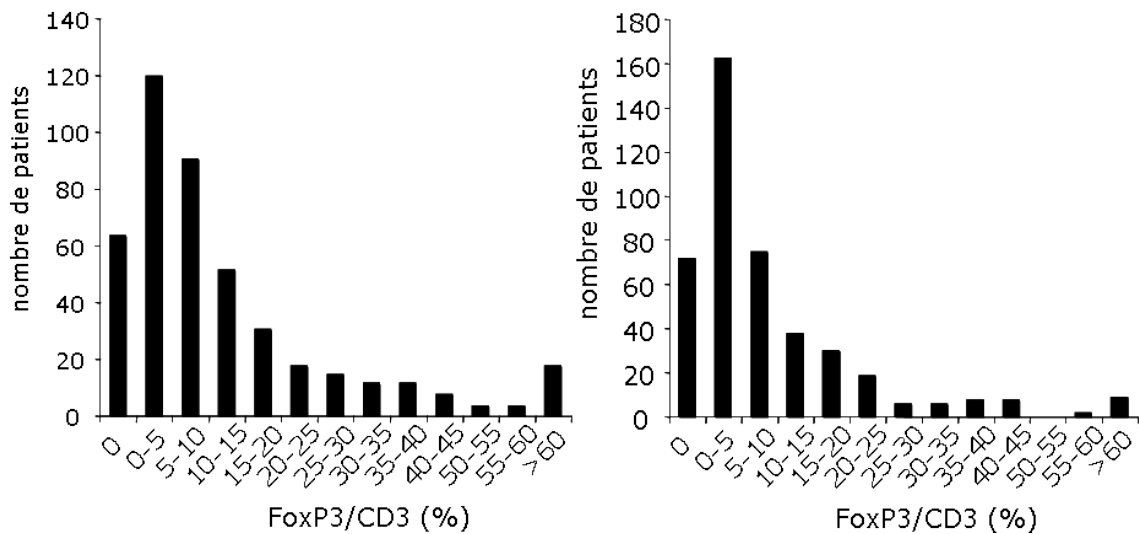


FIGURE 15 – Représentation graphique de la répartition du rapport des densités de cellules FoxP3+ et CD3+ en pourcentage, dans chaque région de la tumeur.

( $p=0,06$ ). Les patients avec un ratio élevé dans les deux régions de la tumeur ont une meilleure survie sans récurrence que les patients avec un faible ratio dans les deux régions. Les patients avec un ratio discordant entre les deux régions de la tumeur ont une survie sans récurrence intermédiaire.

### Les lymphocytes T régulateurs et les lymphocytes Th17

Les lymphocytes Th17 sont de mauvais pronostic (article 4). Les lymphocytes T régulateurs pourraient contrôler l'inflammation induite par les lymphocytes Th17. J'ai fait le rapport densité de cellules FoxP3+ sur celle de cellules IL17+ (figure 17). Tous les cas de figures sont rencontrés : Certains patients n'ont pas de cellules FoxP3+ et des cellules IL17+ (ratio=0), d'autres ont des cellules FoxP3+ et pas de cellules IL17+. Ces derniers cas sont répertoriés dans le groupe  $\text{ratio} \geq 10$ . Trois groupes de patients sont facilement identifiables sur le graphique représentant la répartition du ratio FoxP3+/IL17+ (figure 17) : Ceux avec un ratio inférieur à 0,001 (nommés "lo"), ceux avec un ratio supérieur à 10 ("Hi") et ceux avec un ratio intermédiaire ("Int"). Ces groupes sont valables pour le centre et le front de progression de la tumeur.

Des courbes de survie sans récurrence ont ensuite été tracées à partir de ces groupes (figure 18). Dans le centre de la tumeur, les patients avec un ratio FoxP3+/IL17+ élevé ont une meilleure survie sans récurrence que les patients avec un ratio faible ( $p=0,01$ ). La courbe de survie des patients avec un ratio intermédiaire est comparable à celle des patients avec un ratio élevé. Comme dans le centre de la tumeur, les patients avec un ratio élevé dans le front de progression, ont une meilleure survie sans récurrence que les patients avec un ratio faible ( $p=0,04$ ). Par contre, les patients avec un ratio intermédiaire ont une courbe de survie sans récurrence comparable à celle des patients avec un ratio élevé. La combinaison des deux régions donne des résultats similaires. Le taux de survie sans récurrence des patients à cinq ans est de respectivement 85%, 67% et 40% pour les groupes avec des ratios concordants dans les deux régions fort, intermédiaire, faible. Les patients avec des ratios non concordants ont un taux de survie sans récurrence à cinq ans de 55%.

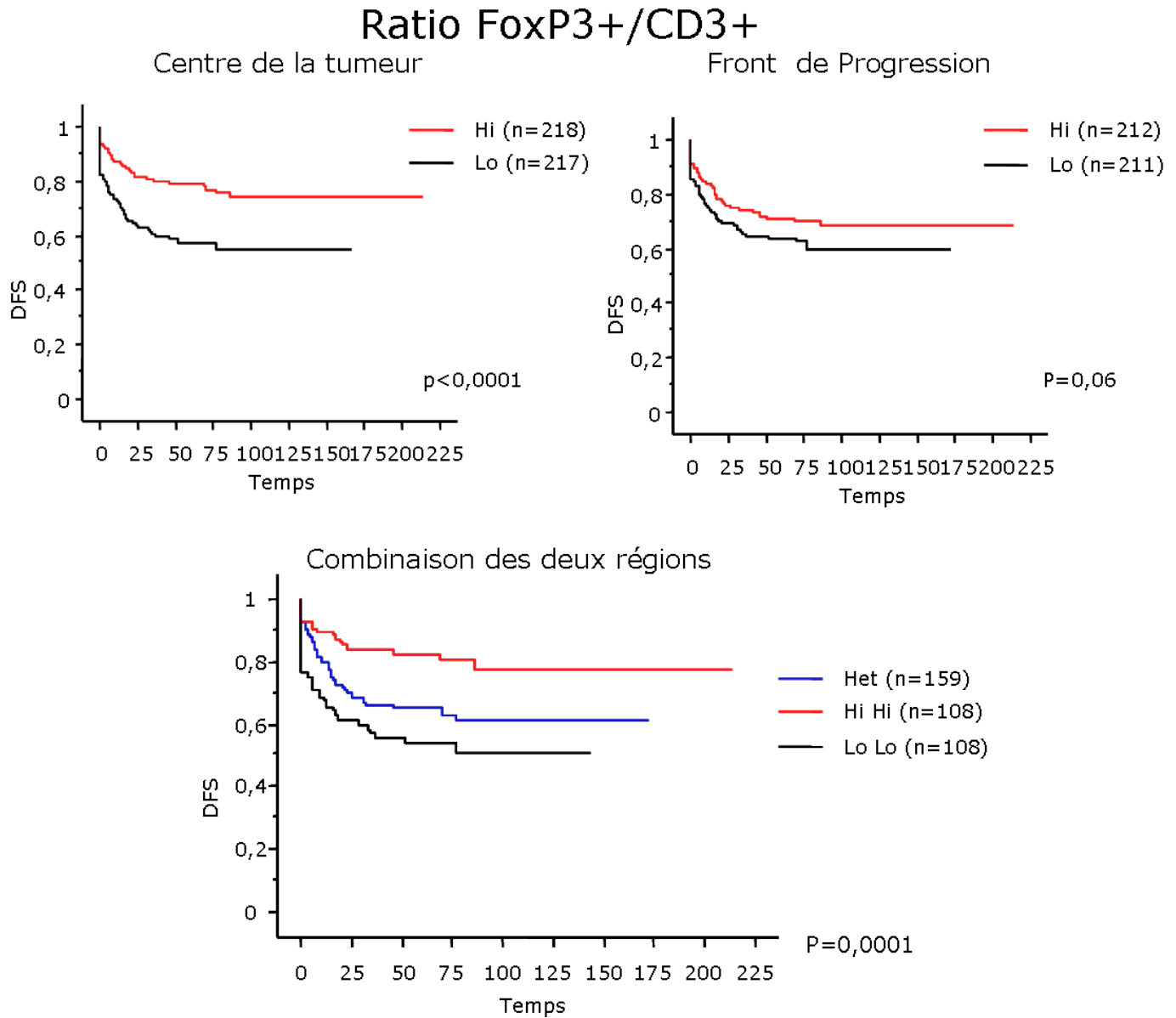


FIGURE 16 – Courbes de survie sans récurrence associées au ratio des densités de cellules FoxP3+ sur celles des cellules CD3+. Pour chaque région de la tumeur, les patients sont classés en groupe avec un fort (Hi, courbe en rouge) ou un faible (Lo, courbe en noir) ratio FoxP3/ CD3 par rapport à la médiane. Ces groupes sont ensuite combinés. Les patients avec un fort ratio (Hi Hi) dans les deux régions sont représentés en rouge, ceux avec un faible ratio (Lo Lo) dans les deux régions en noir et ceux avec un ratio discordant dans les deux régions de la tumeur (Het) en bleu. La survie sans récurrence de ces groupes est visualisée par les courbes de Kaplan-Meier.

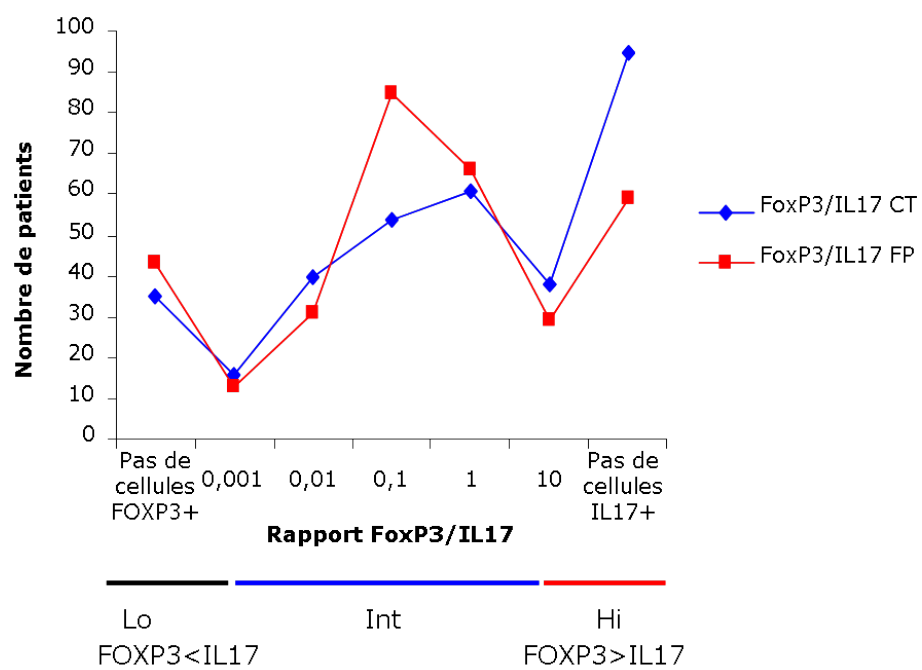


FIGURE 17 – Représentation graphique de la répartition du rapport des densités de cellules FoxP3+ et IL17+, dans le centre (courbe bleu) et le front de progression (courbe rouge) de la tumeur.

## Rapport FoxP3+/IL17+

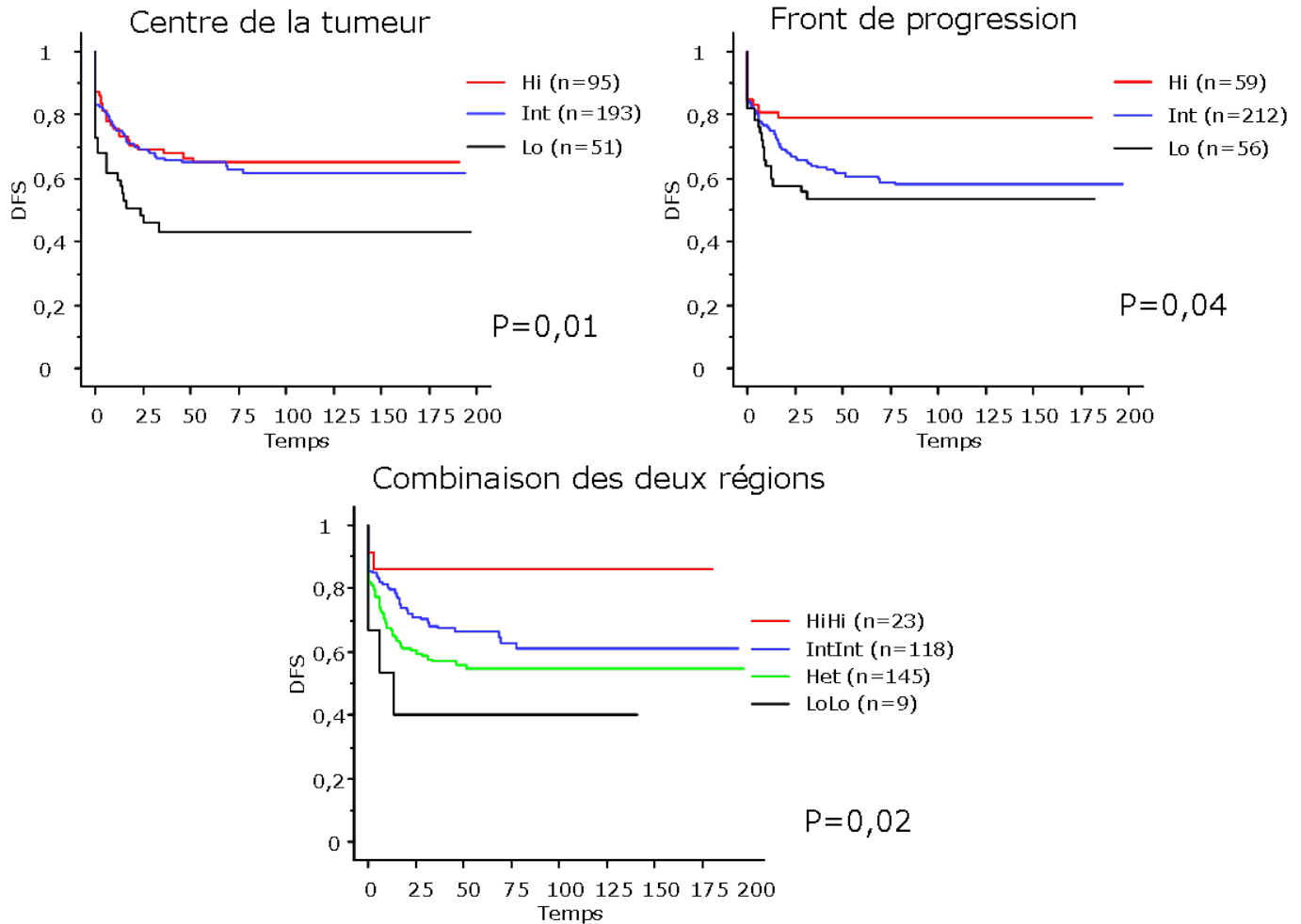


FIGURE 18 – Courbes de survie sans récurrence associées au ratio des densités de cellules FoxP3+ sur celles des cellules IL17+. Pour chaque région de la tumeur, les patients sont classés en groupe avec un ratio fort (Hi, courbe en rouge), intermédiaire (Int, courbe bleu) ou faible (Lo, courbe en noir). Ces groupes sont ensuite combinés. Les patients avec un fort ratio (Hi Hi) dans les deux régions sont représentés en rouge, ceux avec un faible ratio (Lo Lo) dans les deux régions en noir, ceux avec un ratio intermédiaire (Int Int) dans les deux régions en bleu et ceux avec un ratio discordant dans les deux régions de la tumeur (Het) en vert. La survie sans récurrence de ces groupes est visualisée par les courbes de Kaplan-Meier.

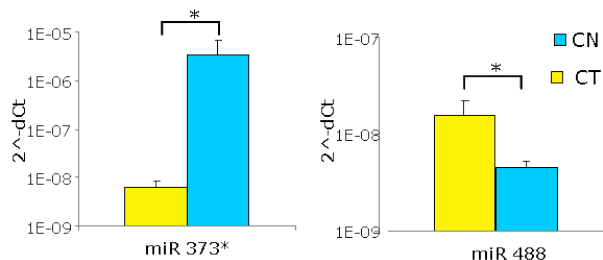


FIGURE 19 – Deux exemples de microARNs différentiellement exprimés sont représentés. Les niveaux d’expression relatifs sont comparés dans les tissus tumoraux (barre jaune) par rapport aux tissus sains (barres bleu).

## Les microARNs

### Matériel et Méthodes

#### PCR en temps réel

J’ai étudié l’expression de 368 microARNs différents sur 104 tumeurs de patients différents ainsi que trois pools de tissu sain à distance de la tumeur par RT-PCR en temps réel. Les ARN totaux ont été extraits à partir de tissus congelés en utilisant la méthode "RNAnow" (Biogentex). L’intégrité et la qualité des ARN ont été évalués sur le BioAnalyser-2100 (Agilent Technologies). Une transcription inverse spécifique pour un pool de 48 microARNs puis la PCR quantitative en temps réel sur des Low-Density-Arrays sur un Taqman 7900HT robotisé (Applied-Biosystems) ont été réalisées conformément aux instructions du fabricant. Les RNU44 et RNU48 ont été utilisés comme contrôle interne. Les données ont été analysées grâce au logiciel SDS v2.3 (Applied-Biosystems).

#### Analyses statistiques

Le logiciel Genesis a été utilisé pour visualiser les matrices de corrélation et pour réaliser le classement hiérarchique "Pearson un-centered". Les courbes de Kaplan-Meier ont été utilisées pour visualiser l’influence des paramètres sur la survie sans récurrence des patients. Le test de Wilcoxon a été utilisé par des comparaisons deux à deux.  $p \leq 0,05$  a été considéré comme statistiquement significatif. Toutes ces analyses ont été faites avec les logiciels statistiques R et Statview.

## Résultats

### Profil d’expression dans les tissus sains et tumoraux

Le profil d’expression des microARNs entre les tissus tumoraux et les tissus sains à distance est différent. Sur les 368 microARNs analysés, 163 sont significativement différents (Tableau 4-5,  $p \leq 0,05$ ). La plupart (132 microARNs sur les 163 significatifs) sont sur-exprimés dans les tissus tumoraux (Tableau 4). Le petit nombre restant (31 microARNs) est sous-exprimé dans les tissus tumoraux (tableau 5). Pour ce dernier groupe, la différence d’expression entre les tissus sains et tumoraux est

miR	P	miR	P	miR	P	miR	P
miR 106b	0,010	miR 202*	0,033	miR 425	0,004	miR 601	0,013
miR 107	0,003	miR 203	0,026	miR 425 5p	0,022	miR 607	0,005
miR 10a	0,012	miR 20a	0,004	miR 429	0,031	miR 615	0,007
miR 128b	0,030	miR 21	0,004	miR 432*	0,030	miR 616	0,006
miR 130b	0,003	miR 210	0,025	miR 449b	0,003	miR 618	0,015
miR 132	0,021	miR 214	0,049	miR 452	0,009	miR 622	0,034
miR 134	0,015	miR 217	0,040	miR 452*	0,006	miR 629	0,043
miR 135a	0,003	miR 22	0,048	miR 484	0,015	miR 639	0,027
miR 135b	0,001	miR 221	0,002	miR 485 5p	0,017	miR 647	0,006
miR 140	0,048	miR 222	0,004	miR 488	0,047	miR 649	0,017
miR 146a	0,018	miR 223	0,026	miR 491	0,002	miR 656	0,004
miR 146b	0,029	miR 224	0,003	miR 500	0,002	miR 660	0,012
miR 148a	0,023	miR 23a	0,013	miR 501	0,005	miR 661	0,029
miR 151	0,005	miR 24	0,009	miR 511	0,045	miR 7	0,017
miR 155	0,024	miR 25	0,007	miR 512 3p	0,029	miR 92	0,002
miR 15a	0,012	miR 29a	0,011	miR 514	0,030	miR 93	0,003
miR 15b	0,023	miR 302a	0,033	miR 517c	0,011	miR 95	0,003
miR 16	0,010	miR 302b	0,030	miR 518b	0,011	RNU6B	0,017
miR 17 3p	0,040	miR 31	0,002	miR 521	0,034		
miR 17 5p	0,005	miR 32	0,047	miR 523	0,020		
miR 181b	0,004	miR 320	0,010	miR 532	0,015		
miR 181c	0,013	miR 325	0,012	miR 548b	0,019		
miR 181d	0,003	miR 326	0,016	miR 548d	0,029		
miR 182	0,002	miR 330	0,011	miR 549	0,027		
miR 183	0,002	miR 331	0,006	miR 550	0,004		
miR 186	0,023	miR 335	0,005	miR 552	0,008		
miR 189	0,023	miR 339	0,013	miR 554	0,019		
miR 18a	0,004	miR 34a	0,002	miR 556	0,047		
miR 191	0,008	miR 361	0,002	miR 558	0,015		
miR 193b	0,008	miR 362	0,011	miR 564	0,006		
miR 196a	0,020	miR 367	0,025	miR 572	0,012		
miR 196b	0,027	miR 374	0,034	miR 576	0,010		
miR 197	0,014	miR 376a*	0,029	miR 579	0,009		
miR 19a	0,008	miR 380 3p	0,006	miR 585	0,040		
miR 19b	0,007	miR 380 5p	0,031	miR 589	0,028		
miR 200a*	0,021	miR 382	0,043	miR 591	0,015		
miR 200b	0,017	miR 409 5p	0,005	miR 593	0,038		
miR 200c	0,019	miR 424	0,006	miR 596	0,031		

TABLE 4 – microARNs significativement exprimés plus fortement dans les tissus tumoraux

miR	p	miR	p	miR	p
miR 1	0,031	miR 133a	0,009	miR 133b	0,045
miR 137	0,007	miR 145	0,045	miR 220	0,012
miR 302a*	0,016	miR 368	0,042	miR 373*	0,018
miR 383	0,005	miR 448	0,005	miR 490	0,015
miR 507	0,004	miR 510	0,038	miR 512 5p	0,009
miR 513	0,005	miR 516 5p	0,005	miR 519b	0,011
miR 524	0,037	miR 526b	0,016	miR 548c	0,019
miR 553	0,006	miR 563	0,026	miR 575	0,017
miR 578	0,013	miR 586	0,006	miR 603	0,011
miR 633	0,004	miR 658	0,007	miR 662	0,005
miR 9	0,015				

TABLE 5 – microARNs significativement exprimés plus faiblement dans les tissus tumoraux

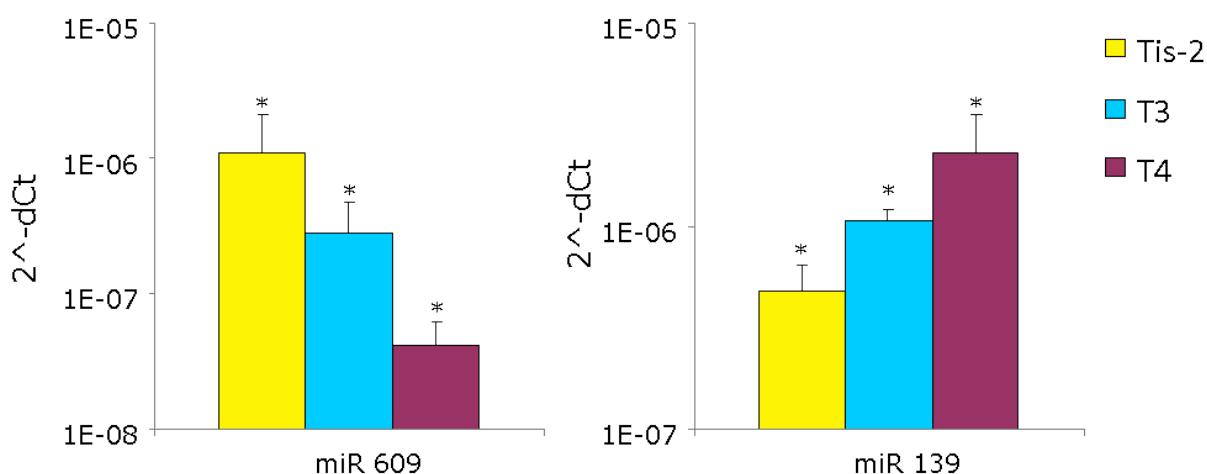


FIGURE 20 – Les niveaux d’expression relatifs sont comparés dans les stades Tis-2 (barre jaune), T1 (barre bleu) et T2 (barre violette). Deux exemples de microARNs différentiellement exprimés sont représentés.

au maximum divisée par 5, à l’exception de miR-373\* dont l’expression est divisée par plus de 500 dans les tissus tumoraux (figure 19). On peut observer des différences bien plus importantes dans le groupe des microARNs sur-exprimés. Ainsi, l’expression du miR-521 est multipliée par plus 50 500 dans les tissus tumoraux. On peut noter également la sur-expression du miR-488 (figure 19).

### Expression des microARNs et progression tumorale

L’expression de certains microARNs évolue avec la progression tumorale (stade T, N et M).

Ainsi l’expression de 41 microARNs est significativement différente entre les différents stades T (figure 6). L’expression de certains microARNs, comme miR-609 diminue progressivement avec l’extension pariétale de la tumeur (figure 20). A l’inverse, l’expression d’autres microARNs augmente progressivement (figure 20). On peut également observer une diminution (ou une augmentation) à un stade particulier (miR-521 par exemple).

miR	p entre Tis-2 vs T3	p entre Tis-2 vs T4	p entre T4 vs T3	miR	p entre Tis-2 vs T3	p entre Tis-2 vs T4	p entre T4 vs T3
miR-125b	0,001	0,024	0,542	miR-34b	0,027	0,140	0,885
miR-198	0,022	0,024	0,877	miR-34c	0,026	0,188	0,991
miR-133a	0,005	0,626	0,027	miR-365	0,045	0,281	0,935
let-7c	0,034	0,446	0,484	miR-381	0,003	0,054	0,491
let-7e	0,037	0,151	0,903	miR-383	0,012	0,281	0,262
miR-1	0,007	0,423	0,094	miR-429	0,045	0,232	0,898
miR-100	0,006	0,188	0,214	miR-433	0,045	0,984	0,198
miR-125a	0,011	0,281	0,602	miR-453	0,038	0,066	0,892
miR-126	0,031	0,299	0,610	miR-484	0,034	0,545	0,690
miR-126*	0,046	0,830	0,292	miR-493	0,039	0,175	0,300
miR-127	0,041	0,682	0,235	miR-497	0,007	0,423	0,235
miR-130a	0,008	0,151	0,405	miR-521	0,045	0,953	0,054
miR-137	0,027	0,401	0,313	miR-650	0,031	0,358	0,326
miR-139	0,011	0,163	0,965	miR-99a	0,011	0,093	0,571
miR-143	0,005	0,110	0,587	miR-99b	0,014	0,119	0,956
miR-145	0,007	0,175	0,334	miR-9*	0,005	0,072	0,579
miR-195	0,027	0,711	0,231	miR-378	0,445	0,032	0,045
miR-199a	0,015	0,216	0,424	miR-549	0,996	0,045	0,009
miR-199b	0,023	0,264	0,347	miR-515-5p	0,420	0,049	0,120
miR-200a*	0,017	0,062	0,698	miR-580	0,083	0,045	0,610
miR-214	0,035	0,129	0,972	miR-133b	0,061	0,520	0,020
miR-218	0,016	0,078	0,868	miR-146a	0,936	0,093	0,020
miR-296	0,047	0,163	0,799	miR-148a	0,894	0,078	0,035
miR-30a-3p	0,041	0,800	0,253	miR-302c*	0,560	0,281	0,050
miR-342	0,012	0,599	0,074	miR-609	0,575	0,175	0,050

TABLE 6 – microARNs significativement différentiellement exprimés entre les différents stades T ( $p \leq 0,05$ )

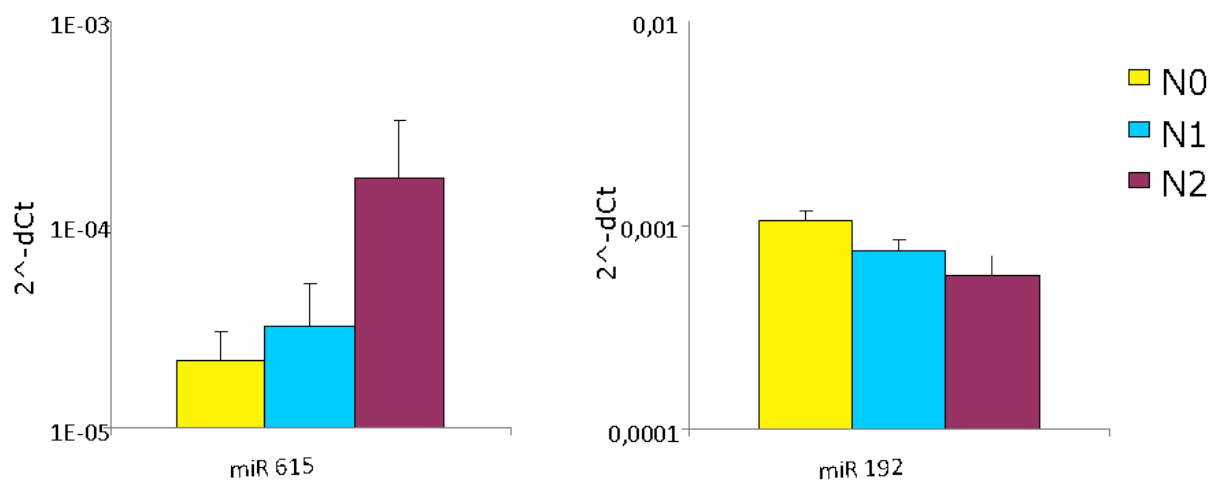


FIGURE 21 – Les niveaux d'expression relatifs sont comparés dans les stades N0 (barre jaune), N1 (barre bleu) et N2 (barre violette). Deux exemples de microARNs différentiellement exprimés sont représentés.



miR	p entre N0 vs N1	p entre N0 vs N2	p entre N1 vs N2	miR	p entre N0 vs N1	p entre N0 vs N2	p entre N1 vs N2
miR-181b	0,011	0,404	0,010	let-7e	0,131	0,112	0,018
miR-181d	0,042	0,616	0,025	miR-103	0,261	0,112	0,022
miR-30e-3p	0,042	0,532	0,044	miR-106b	0,138	0,350	0,021
miR-34a	0,018	0,906	0,026	miR-125a	0,216	0,061	0,007
miR-369-3p	0,007	0,697	0,016	miR-126*	0,322	0,188	0,048
miR-345	0,016	0,743	0,070	miR-130a	0,579	0,084	0,024
miR-373	0,029	0,771	0,054	miR-132	0,486	0,172	0,031
miR-651	0,040	0,469	0,275	miR-140	0,286	0,053	0,009
miR-125b	0,480	0,024	0,007	miR-145	0,291	0,180	0,034
miR-126	0,694	0,007	0,004	miR-146b	0,375	0,308	0,042
miR-139	0,839	0,049	0,036	miR-183	0,082	0,218	0,044
miR-143	0,346	0,042	0,004	miR-191	0,057	0,439	0,017
miR-152	0,291	0,015	0,002	miR-198	0,242	0,087	0,032
miR-16	0,573	0,043	0,013	miR-21	0,253	0,268	0,046
miR-187	0,904	0,021	0,029	miR-218	0,322	0,096	0,032
miR-193a	0,682	0,046	0,024	miR-23a	0,186	0,397	0,048
miR-199a	0,619	0,015	0,047	miR-24	0,138	0,094	0,008
miR-199a*	0,693	0,022	0,034	miR-26a	0,196	0,196	0,034
miR-199b	0,731	0,018	0,017	miR-299-5p	0,146	0,268	0,040
miR-214	0,976	0,032	0,016	miR-30a-3p	0,407	0,115	0,036
miR-22	0,365	0,028	0,003	miR-30b	0,135	0,532	0,042
miR-365	0,984	0,019	0,016	miR-330	0,159	0,325	0,036
miR-372	0,535	0,050	0,029	miR-331	0,704	0,086	0,024
miR-489	0,729	0,043	0,025	miR-34b	0,423	0,253	0,044
miR-653	0,659	0,005	0,021	miR-34c	0,112	0,164	0,012
miR-105	0,174	0,016	0,449	miR-376a	0,299	0,082	0,013
miR-192	0,102	0,003	0,107	miR-411	0,308	0,214	0,012
miR-194	0,119	0,005	0,218	miR-425	0,095	0,673	0,044
miR-200a*	0,069	0,018	0,511	miR-484	0,619	0,118	0,040
miR-200b	0,171	0,007	0,164	miR-487b	0,336	0,160	0,011
miR-200c	0,068	0,039	0,631	miR-501	0,317	0,099	0,028
miR-215	0,220	0,038	0,415	miR-517c	0,156	0,274	0,034
miR-378	0,202	0,044	0,536	miR-520g	0,110	0,279	0,026
miR-412	0,154	0,043	0,426	miR-615	0,765	0,130	0,040
miR-493	0,128	0,032	0,540	miR-99a	0,666	0,059	0,025
miR-520b	0,249	0,047	0,283	miR-99b	0,694	0,061	0,013
let-7c	0,286	0,130	0,048				

TABLE 7 – microARNs significativement différentiellement exprimés entre les différents stades N.(p≤0,05)

miR	p	miR	p
miR 196b	0,04	miR 523	0,03
miR 216	0,03	miR 526b	0,008
miR 302b	0,03	miR 548c	0,03
miR 331	0,04	miR 551b	0,03
miR 373etoile	0,02	miR 555	0,009
miR 500	0,01	miR 601	0,02
miR 501	0,02	miR 661	0,01
miR 503	0,01	RNU6B	0,01
miR 515 3p	0,03		

TABLE 8 – microARNs significativement différenciellement exprimés entre les différents stades M. ( $p \leq 0,05$ )

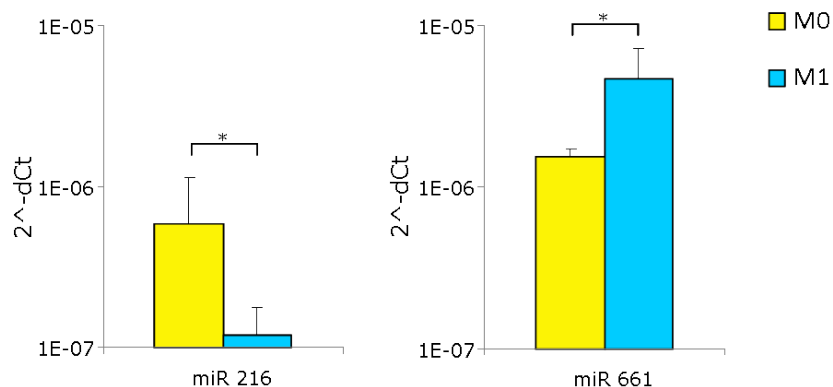


FIGURE 22 – Les niveaux d’expression relatifs sont comparés dans les stades M0 (barre jaune) et M1 (barre bleu). Deux exemples de microARNs différenciellement exprimés sont représentés.

De la même manière, l’expression de 71 microARNs est différente entre les stades N (figure 7). Ainsi l’expression du miR-615 augmente progressivement avec l’envahissement des ganglions lymphatiques tandis que l’expression du miR-192 diminue (figure 21). L’expression d’un microARN peut également augmenter (ou diminuer) à un stade particulier (miR-369 3p, par exemple).

Entre les stades M, seul un petit nombre (17) de microARNs est différenciellement exprimé (Tableau 8). L’expression des microARNs peuvent augmenter avec la présence de métastases, comme miR-661, ou, au contraire diminuer, comme miR-216 (figure 22).

## Expression des microARNs et récurrence

Afin d'identifier les microARNs associés à la récurrence des patients, les courbes de survie sans récurrence ont été tracées. Les patients ont été séparés en deux groupes (fort, courbe rouge, et faible, courbe noire) suivant le niveau d'expression du microARN par rapport à la médiane de tous les patients. 65 microARNs se sont révélés significatifs. Leurs "hazard ratio" sont représentés en figure 23. Les microARNs peuvent être séparés en deux groupes : Ceux qui sont de bon pronostic ( $HR \leq 1$ ) comme le miR-488 et ceux qui sont de mauvais pronostic ( $HR \geq 1$ ) comme le miR-660. En effet, le taux de survie sans récurrence à 5 ans des patients ayant une forte expression du miR-488 est de 78% contre 38% pour les patients avec une faible expression. Au contraire, le taux de survie sans récurrence à 5 ans des patients ayant une faible expression du miR-660 est de 77% contre 40% pour les patients avec une forte expression (figure 23). En combinant l'expression de ces deux microARNs, on obtient un taux de survie de 80% pour le groupe de patients avec une forte expression de miR-488 et une faible expression de miR-660 (courbe rouge), de seulement 30% pour le groupe de patients avec une faible expression de miR-488 et une forte expression de miR-660 et de 70% pour une autre combinaison de ces microARNs.

## microARNs et système immunitaire

Les données relatives aux microARNs ont été mises en relation avec les données sur les ARNm et sur les protéines *in situ* du système immunitaire. Pour cela, j'ai réalisé une matrice de corrélation entre ces différentes données (données non montrées). Je me suis plus particulièrement intéressée aux trois microARNs significatifs pour la survie sans récurrence : miR-657 et miR-488, de bon pronostic et miR-660 de mauvais pronostic.

Des extraits sur les zones correspondantes à ces microARNs ont été réalisés sur la matrice du système immunitaire inné et celle du système immunitaire adaptatif (figure 24 et 25).

Les miR-657 et miR-488 ne présentent pas de corrélation négative avec les données du système immunitaire inné (figure 24). Par contre, miR-660 a une légère corrélation négative avec les macrophages (CD68) au centre de la tumeur (CT) et, en moyenne dans la tumeur (CT-IM). Il a également une légère corrélation négative avec les neutrophiles (granulocytes) quelque soit la région de la tumeur (figure 24).

Au niveau du système immunitaire adaptatif (figure 25), miR-657 présente une corrélation positive avec la présence de lymphocytes T (CD3) et T mémoires (CD45RO), ainsi qu'avec la présence d'IL23 au niveau protéique. Le miR-488 présente les mêmes corrélations positives, avec, en plus, une corrélation positive avec les ARNm codant pour les récepteurs CCR3 et CCR4. Ces deux microARNs ne présentent pas de corrélation négative avec les marqueurs étudiés. Au contraire, miR-660 présente de nombreuses corrélations négatives. Il corrèle négativement avec des marqueurs Th1 au niveau ARNm (IL18, IFN $\gamma$ , IRF1, STAT1, TAP1, CD38, IL27, CCL5, CCR5, CD8A et TBX21) et protéique (CD8), avec des marqueurs de cytotoxicité au niveau des ARNm (PRF1 et GNLY) et protéique (GZMB et CD57) et quelques ARNm de marqueurs de lymphocytes T régulateur (IL10 et CTLA4). On peut noter la légère corrélation positive avec les lymphocytes T (CD3 *in situ*). Ces données suggèrent une action de miR-660 qui aboutit à l'inhibition de la réponse Th1/cytotoxique des lymphocytes T.

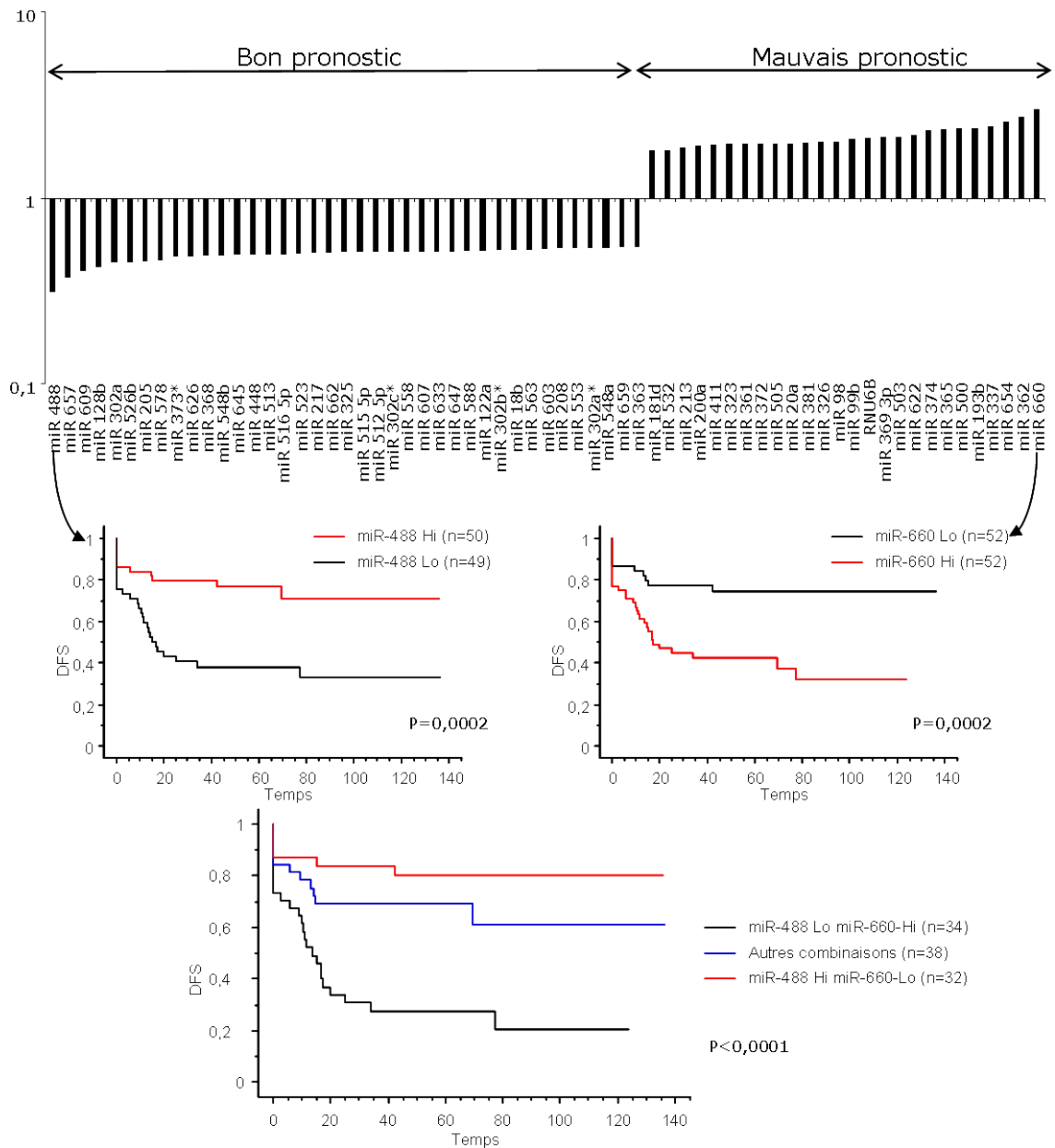


FIGURE 23 – (A) Les "Hazard ratio" (HR) des comparaisons des courbes de survie sans récurrence de chaque marqueurs sont représentés. Les patients sont séparés en deux groupes par rapport à la médiane de l'expression du microARN observé. (B) Les courbes de survie sans récurrence du miR-488 et miR-660 sont représentés, seuls ou en combinaison.

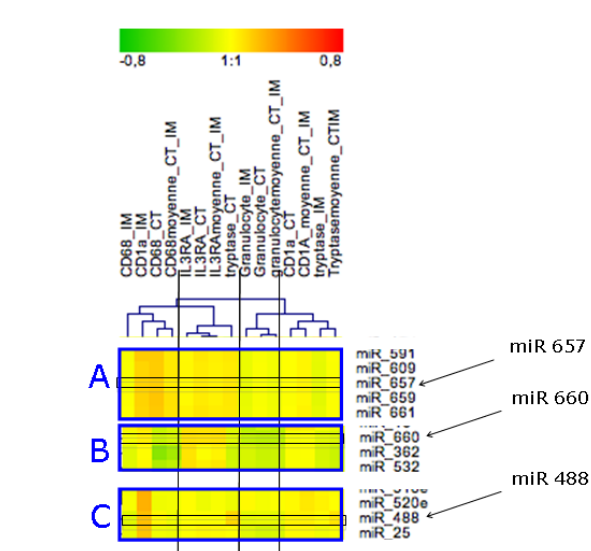


FIGURE 24 – Extrait de la matrice de corrélation entre les données sur les microARNs et les cellules immunitaires innées. Les parties A, B et C sont tirés de la matrice de corrélation entre les microARNs et les cellules immunitaires innées. Elles contiennent respectivement les corrélations avec miR-657, miR-488 et miR-660.

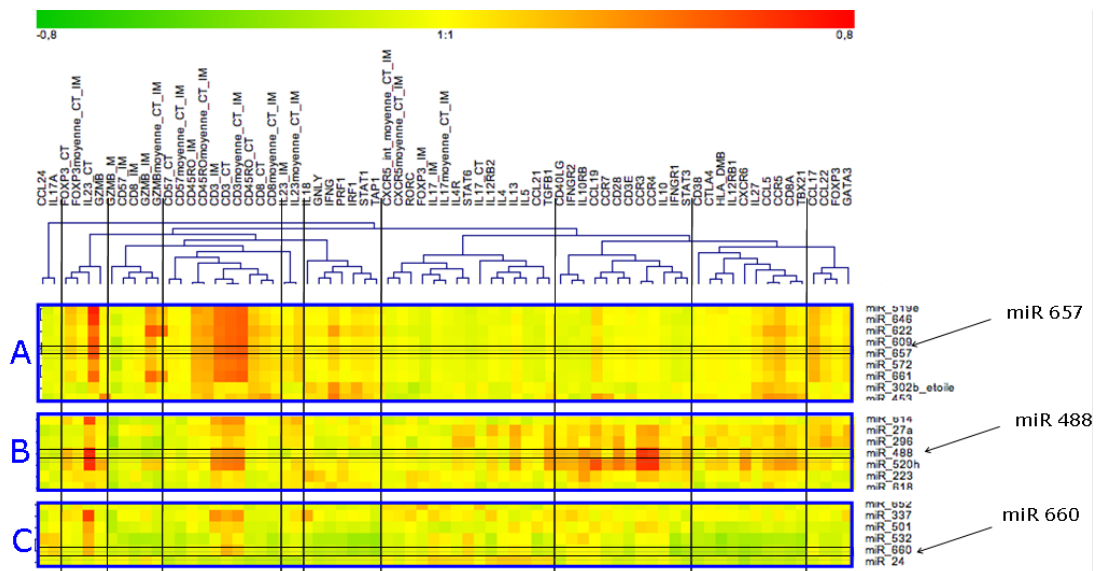


FIGURE 25 – Extrait de la matrice de corrélation entre les données sur les microARNs et les cellules immunitaires adaptatives. Les parties A, B et C sont tirés de la matrice de corrélation entre les microARNs et les cellules immunitaires innées. Elles contiennent respectivement les corrélations avec miR-657, miR-660 et miR-488.

# Discussion

# La réaction immunitaire dans les cancers colorectaux

Le cancer colorectal est une maladie complexe. Actuellement, les critères anatomopathologiques déterminent le traitement des patients, mais ils ne permettent pas une identification suffisamment précise des patients à risques. Par conséquent, certains patients ne bénéficient pas d'un traitement efficace. Or depuis quelques décennies, les recherches ont démontré que le système immunitaire jouait un rôle important dans le contrôle de la progression tumorale chez l'Homme dans de nombreux cancers.

Nous avons montré que le type, la densité et la localisation des cellules immunitaires, permettent de mieux prédire la survie des patients que les données histologiques utilisées actuellement (article 1). En effet, l'analyse par immunohistochimie en Tissue-MicroArray (TMA) dans deux zones tumorales (centre et front de progression de la tumeur) des marqueurs lymphocytaires CD3 et CD45RO ont une valeur pronostique puissante en terme de survie globale ( $p < 10^{-7}$ ) et de survie sans récurrence ( $p < 10^{-8}$ ) des patients bien supérieure à celle des paramètres anapathologiques (stade I, II et III). Ces observations ont été confirmées de façon indépendante sur deux autres cohortes de patients. L'analyse combinatoire des densités de cellules CD3+ et des cellules CD45RO+ a permis l'identification précise et sans ambiguïté des patients à risques élevés de récurrence ( $p < 10^{-4}$ ) indépendamment des paramètres anapathologiques. Cette étude montre pour la première fois le caractère pronostic puissant des paramètres immunologiques (type, densité et localisation des cellules immunitaires) chez les patients atteints de cancers colorectaux. En tenant compte de ces paramètres, les critères d'invasion tumorale (TNM) ne sont plus significatifs et dépendent même des paramètres immunitaires. Ces résultats ont été confirmés par une étude complémentaire (annexe 2) portant sur les tumeurs de petite taille, sans envahissement ganglionnaire ni métastase (stade I et II).

La détection de métastases reste un élément central dans la prise de décision de l'administration de traitements adjuvants aux patients atteints de cancers colorectaux. Nous avons étudié les paramètres cellulaires et moléculaires qui pourraient décrire les réponses immunitaires *in situ* dans des cancers colorectaux humains en fonction de paramètres cliniques tel que l'envahissement de ganglions ou d'organes à distance (META- ou META+). L'analyse des corrélations du phénotypage effectuée par cytométrie de flux de tumeurs sans métastases (META-) et avec un fort infiltrat lymphocytaire, montre différents clusters. Un cluster est associé à des cellules immunitaires innées (macrophages, cellules dendritiques, natural killer) et des cellules T en phase précoce d'activation. Un autre cluster est associé aux marqueurs de cytotoxicité et à la sous-population de cellules T effectrices mémoires. Ce profil est altéré dans les tumeurs avec un faible infiltrat et absent dans des tumeurs avec métastases (META+). La co-expression du gène granulysine (cytotoxicité) et IRF1 (Th1) permet de prédire la survie du

patient indépendamment du statut métastatique. Cependant l'expression du VEGF (vascularisation) réduit significativement la survie des patients exprimant de forts niveaux de granulysine.

Nous avons pu ainsi déterminer quatre groupes principaux de patients. Le premier groupe présente une forte densité de cellules T avec une bonne coordination immunitaire et un niveau faible de VEGF. Son risque de récurrence à 2 et 10 ans est identique et de 10%. Le deuxième et le troisième groupe ont une forte densité de cellules T mais soit une perte de la coordination de la réponse immunitaire, soit un niveau élevé de VEGF. Leur risque de récurrence à 2 ans est alors de 50%. Le quatrième groupe a le risque de récurrence sur 2 ans le plus élevé (80%). Ces patients ont des densités de lymphocytes T bas et pas de coordination de la réponse immunitaire.

Nous avons vu l'importance de la réponse immunitaire dans le contrôle du développement de la tumeur. Or les lymphocytes T infiltrant la tumeur sont la résultante de diverses variables tumorales. Le niveau d'expression HLA, la présence d'antigènes tumoraux, les diverses mutations (instabilité chromosomique, déficiences des gènes de réparation,...) et les facteurs présents dans le microenvironnement (VEGF par exemple) ont une influence sur la réponse immunitaire. La réponse immunitaire au site primaire de la tumeur permet d'évaluer le risque d'une récurrence post-chirurgicale. Or, la tumeur primaire est enlevée lors de la chirurgie. L'infiltrat immunitaire se trouvant dans la tumeur enlevée, doit donc refléter une qualité, une quantité de cellules immunitaires se trouvant en systémique. Ces cellules permettraient d'éliminer les cellules cancéreuses restantes dans l'organisme.

Divers outils bio-informatiques (article 3 et annexe 1) nous ont permis d'intégrer un grand nombre de données et de construire des réseaux moléculaires. Ces outils performants permettent de créer des hypothèses, vérifiées ensuite, sur les mécanismes sous-jacents les réponses immunitaires pertinentes pour le pronostic des patients. Cette approche nous a révélé l'importance de trois chimiokines (CX3CL1, CXCL10 et CXCL9) ainsi que de trois molécules d'adhésions (ICAM1, VCAM1 et MAD-CAM1) dans l'attraction et donc la présence d'une forte densité de cellules immunitaires. En effet, CXCL10 et CXCL9, contrairement à CX3CL1, corrélaient avec la présence de lymphocytes T CD8 mémoires (CD3+ CD8+ CD45RO+). Les tumeurs de patients ayant une expression élevée de CX3CL1 ont une densité significativement plus élevée de cellules cytotoxiques ( $P \leq 0.05$  pour granzyme B et CD57 au centre et au front de progression de la tumeur). CX3CL1 semble attirer également une sous-population T avec des TCR particuliers. En effet, en présence d'une forte expression de CX3CL1, les niveaux récepteurs  $V\beta L03$ ,  $V\beta L07$ ,  $V\beta L06$  et  $V\beta L09$  sont augmentés ( $P \leq 0.01$ ,  $P \leq 0.008$ ,  $P \leq 0.05$  et  $P \leq 0.05$ , respectivement). Ces TCR particuliers corrélaient avec le bon pronostic des patients. La stratégie permettant l'intégration des données, une reconstruction des réseaux moléculaires et une création d'hypothèses est une stratégie puissante pour découvrir les mécanismes moléculaires sous-jacents. Cela nous a permis de démontrer l'importance de chimiokines et molécules d'adhésion pour la présence de fortes densités de lymphocytes T au site de la tumeur.

Une réponse immunitaire de type Th1 et cytotoxique ainsi qu'une bonne coordination est nécessaire au contrôle du développement de la tumeur. Nous avons ensuite étudié la balance entre les cellules cytotoxiques et les sous-populations de lymphocytes T helper ainsi que leurs impacts sur la survie sans récurrence des patients. La matrice de corrélation classée de manière non supervisée nous a révélé des groupes fonctionnels de gènes dont nous avons analysé les impacts sur le devenir clinique des patients. Le groupe de patients avec une forte expression des gènes Th1/cytotoxique ont un taux de survie sans récurrence augmenté par rapport aux patients avec une faible expression. Au contraire, les patients avec une faible expression des gènes Th17 semblent être de bons pronostics. L'analyse combinée des Th1/cytotoxiques et des Th17 permet d'avoir une meilleure discrimination de la récurrence.



Aucun patient Th1/cytotoxique fort et Th17 faible ne récidive. A l'opposé, la combinaison la pire en terme de récurrence est Th1/cytotoxique faible et Th17 fort. Ces résultats ont été confirmés par une analyse des protéines *in situ* par immunohistochimie. Les patients ayant une forte densité de cellules CD8+ (marqueur Th1) dans les deux zones de la tumeur, quelque soit la densité de cellules IL17+ (marqueur Th17) n'ont pas fait de récurrence. A l'opposé, les patients avec une faible densité de cellules CD8+ dans les deux zones de la tumeur sont de mauvais pronostic. Les patients avec une densité non concordante dans les deux zones de la tumeur (hétérogènes), peuvent être séparés grâce à la densité de cellules IL17+. Aucun patient CD8 hétérogène et IL17 faible ne récidive tandis que 80% des patients CD8 hétérogènes IL17 fort récidivent dans les deux ans. Les clusters fonctionnels associés aux lymphocytes Th1/cytotoxiques et Th17 ont des effets opposés sur la survie sans récurrence des patients. Leurs informations sont complémentaires bien que l'effet des Th1/cytotoxiques soit supérieur à celles des Th17 sur le pronostic des patients.

Les lymphocytes T sont présents en forte quantité dans une tumeur de petite taille puis diminuent progressivement lors de la progression tumorale. Le nombre élevé de lymphocytes Th1/cytotoxiques et le petit nombre de lymphocytes Th17 participent au contrôle du développement de la tumeur. Après la chirurgie, la forte densité de lymphocytes Th1/cytotoxiques en systémique participe à la non-récurrence du patient (figure 26A). Au contraire, la faible densité de lymphocytes Th1/cytotoxiques et la forte densité de lymphocytes Th17 ne contrôlent plus le développement de la tumeur ni la récurrence (figure 26B).

Aucune association avec un pronostic particulier n'a été trouvée avec les clusters comprenant des gènes de lymphocytes T régulateurs (article 4). Or l'étude de la densité de cellules FoxP3+ *in situ* révèle une association avec un bon pronostic des patients (données non publiées). Une étude seule de l'expression de FoxP3 est concordante avec ce résultat (article 4). De plus, cet effet sur le pronostic des patients n'est pas le reflet d'une augmentation parallèle de la densité de cellules FoxP3+ et de lymphocytes T CD3+. En effet, un fort ratio FoxP3/CD3 est de bon pronostic. Dans les cancers colorectaux, l'effet de cellules FoxP3+ sur la récurrence des patients n'est pas unanime. En effet, Loddenkemper et al [142] n'a pas trouvé de différence entre les patients fortement et faiblement infiltrés par les cellules FoxP3+ tandis que Salama et al [143] ont trouvé une survie augmentée des patients avec une forte infiltration de cellules FoxP3+.

Il existe deux clusters clairement dissociés (article 4, figure 2) comprenant des gènes de lymphocytes T régulateurs. Le premier comprenant les cytokines (IL10 et TGF $\beta$ ), le second des marqueurs cellulaires (CTLA4 et FOXP3). Les autres types cellulaires (dont les cellules tumorales [144]) qui ont la capacité de sécréter les cytokines IL10 et TGF $\beta$  peuvent expliquer cette dissociation. Le deuxième cluster, comprenant CTLA4 et FOXP3, regroupe également deux chimiokines, CCL17 et CCL22, impliquées à la fois dans les réponses immunes de type Th2 et Treg. Or ces chimiokines ne sont pas en elle-même de bon pronostic (article 4 figure 2B). L'effet global de ce cluster, comprenant plusieurs gènes ciblant différentes fonctions et différentes cellules, peut différer de l'effet d'une population particulière avec une fonction particulière. Ceci pourrait expliquer l'apparente contradiction entre l'étude de l'expression de gènes et celle des protéines *in situ*.

La réponse immunitaire surtout de type Th1, cytotoxique et mémoire permet de contrôler le développement de la tumeur. Or les lymphocytes T régulateurs ont une fonction de modulation de la réponse immunitaire. Le bon pronostic d'une forte densité de cellules FoxP3+ semble contradictoire. D'un autre côté, l'inflammation a un effet promoteur sur le développement de la tumeur [145]. Le rôle positif de la présence de cellules FoxP3+ pourrait s'expliquer par la nécessité de contrôler l'inflam-

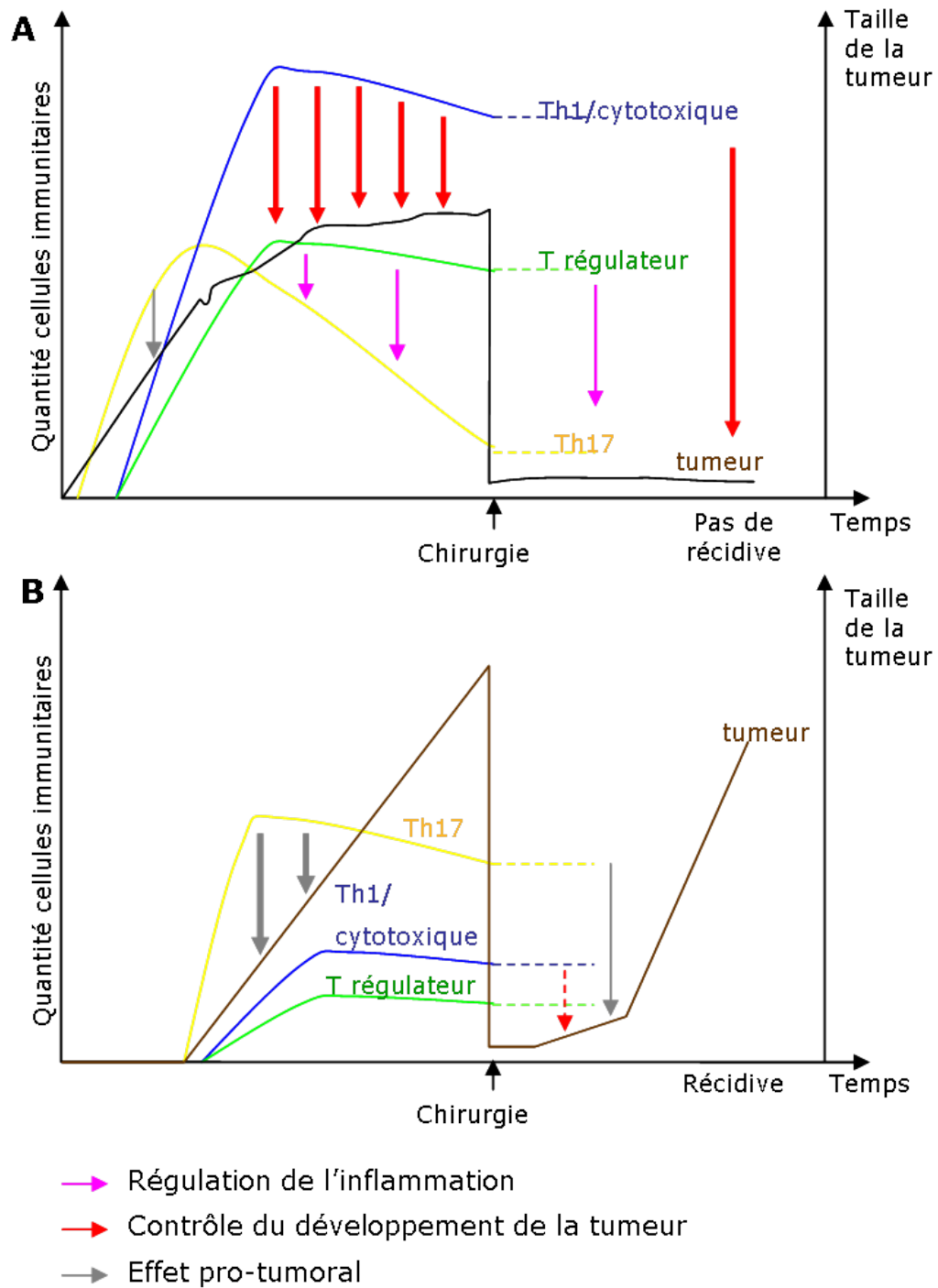


FIGURE 26 – Modèle de l'interaction lymphocyte T / tumeur. (A) Situation idéale d'interaction : Une forte densité de lymphocytes Th1 et cytotoxiques et une faible densité de lymphocytes Th17 contrôlent le développement de la tumeur et la récurrence. (B) Situation la plus péjorative : Une faible densité de lymphocytes Th1 et cytotoxiques et une forte densité de Th17 ne contrôlent pas le développement de la tumeur ni la récurrence.

mation. Leurs présences pourrait être le reflet de ce contrôle passé ou présent (figure 26).

# Les microARNs

Depuis le début de la biologie moléculaire, il était acquis que l'expression d'un gène se faisait de façon linéaire : Le gène, sous le contrôle de différentes protéines régulatrices, semblait être tout d'abord transcrit en un ARN messager qui lui-même était traduit en protéines. Or, de petites séquences d'ARNs découverts récemment, appelées microARN, ont une grande importance dans le contrôle de ce processus. En effet, ils interviennent dans la régulation des ARN messagers et semblent être indispensables au bon fonctionnement des cellules. Ils jouent un rôle dans de nombreux processus biologiques physiologiques (dont le fonctionnement des cellules du système immunitaire) et physio-pathologiques (dont le cancer).

Comme attendu, nous avons trouvé un profil altéré de l'expression des microARNs dans les tissus tumoraux. Les microARNs miR-21, miR-29a, miR-19a, miR-92, miR-107, miR-148, miR-200b, miR-200c, miR-221 et miR-223 sont décrits comme sur-exprimés [125]. Les microARNs miR106b, miR-133, miR-146b, miR-372 et miR-373 qui ont décrits comme oncomiRs dans d'autres types de cancer [113] [112] [114] ont leurs expressions qui augmentent avec la progression tumorale (stade T pour miR-133, stade N pour les autres).

Cependant parmi les microARNs décrits comme étant sous-exprimés dans les cancers colorectaux, seul miR-145 diminue dans notre cohorte. De même, les microARNs décrits comme suppresseurs de tumeur dans d'autres types de cancer, n'ont pas leurs expressions diminuées dans notre cohorte.

Des microARNs ont également été montrés comme prédictifs de la survie. Ainsi, miR-20a qui est associé à un mauvais pronostic dans les cancers colorectaux [125], l'est également dans notre cohorte. Aucun microARN prédictif de la survie dans notre cohorte, n'a été décrit dans les cellules immunitaires. Cependant, miR-214, miR-21, miR-155 et miR-24 dont les expressions augmentent avec l'activation des lymphocytes T CD4+ et CD8+ sont sur-exprimés dans les tissus tumoraux de notre cohorte.

Les différentes sous-populations de lymphocytes T ont un effet différent sur le contrôle du développement tumoral. En effet, les Th17 ont un effet opposé à celui des lymphocytes Th1 ou des lymphocytes T cytotoxiques. Dans l'état actuel des connaissances sur les microARNs, aucun microARN n'est associé à une sous-population de lymphocytes T en particulier. De plus, nous avons étudié l'expression des microARNs sur l'ensemble des cellules comprises dans la masse tumorale. Cela comprend à la fois des cellules tumorales, des cellules immunitaires, des cellules des vaisseaux sanguins et lymphatiques... Le profil d'expression des microARNs observé est la résultante de toutes ces variables. Il est donc intéressant de corrélérer l'expression des microARNs avec l'expression de gènes et la densité de cellules immunitaires. Nous nous sommes plus particulièrement intéressés à des microARNs, miR-488, miR-657 et miR-660 prédictifs de la survie des patients. Nous avons peu de connaissance sur ces trois microARNs. miR-660 est sur-exprimé par la lignée HL-60 (cellules leucémiques) traité

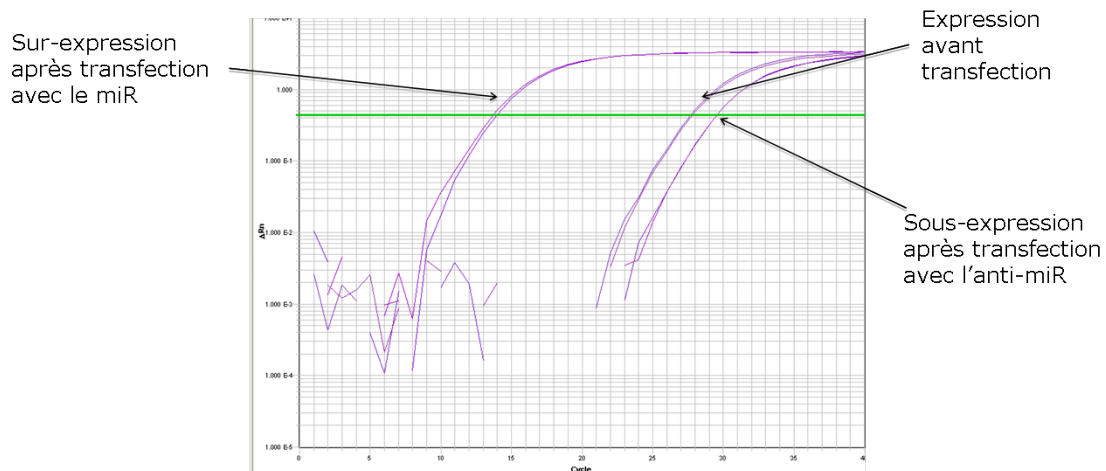


FIGURE 27 – Transfection de cellules pour une sur-expression de miR-657. Les cellules transfectées expriment 15 000 fois plus le miR-657.

par HNE (4-hydroxynonanal). Le HNE est un produit d'oxydation des lipides ayant des propriétés anti-prolifératives et de différenciation cellulaire [146]. La cible décrite de miR-657 est IGF2R (Insuline-like Growth Factor 2 Receptor), associé au diabète de type 2 [147].

miR-488 et miR-657 corrént positivement avec la présence de lymphocytes T CD3+ et T mémoires CD45RO+. Ils ne présentent aucune corrélation négative avec les marqueurs étudiés. miR-660 a un profil différent. Il présente de nombreuses corrélatons négatives avec les marqueurs d'une réaction immunitaire de type Th1 et cytotoxique et une légère corrélation positive avec la densité de lymphocytes T CD3+. Il existe de nombreuses possibilités d'action de ces microARNs. En effet, ces microARNs peuvent être exprimés principalement dans les cellules tumorales et avoir une conséquence sur le système immunitaire. Par exemple, une meilleure présentation antigénique des cellules tumorales aboutirait à une réaction immunitaire plus importante. Les microARNs peuvent également être exprimés par les cellules immunitaires. Ainsi, la présence d'un microARN peut empêcher la production de molécules cytotoxiques, les cellules immunitaires ne pouvant plus détruire les cellules tumorales.

Les cibles, les cellules affectées et les effets de ces microARNs doivent être identifiés. Dans ce but, une analyse *in silico* des cibles et du rôle potentiel des microARNs a été réalisée. Par exemple, le microARN miR-657 a des gènes-cibles potentiels impliqués à la fois dans le système immunitaire et dans les cellules tumorales (tableau 9). Les cibles seront validées par une transfection de diverses lignées cellulaires pour sur-exprimer les microARNs (figure 27). Une étude de l'expression des gènes-cibles potentiels ainsi qu'une étude fonctionnelle seront ensuite réalisées.

GOTerm	Associated Genes Found
activated T cell proliferation	ITGAM
activation of immune response	C2, SERPING1, TLR6
activation of innate immune response	TLR6
adaptive immune response	C2, CD27, EBI3, PRKCD, SERPING1, TLR6
adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	C2, CD27, EBI3, PRKCD, SERPING1, TLR6 C2, CD27, EBI3, PRKCD, SERPING1, TLR6
B cell activation	PRKCD
B cell mediated immunity	C2, CD27, PRKCD, SERPING1
B cell proliferation	PRKCD
cell activation during immune response	TLR6
colorectal	CTNNB1
Colorectal cancer	BAX, CTNNB1, MAPK3
gastric	ERBB2
immune effector process	C2, CD27, DBH, NLRP3, PRDX1, PRDX2, PRKCD, SERPING1
immune response	APLN, ARHGDIB, C1QBP, C2, CCL26, CD27, CSF2, CYBA, CYP27B1, DBH, EBI3, IL1F7, IL1RAPL1, IL20RB, NCF4, NLRP3, PLUNC, PRDX1, PRDX2, PRKCD, SERPING1, TLR6
immune response-activating signal transduction	TLR6
immune system development	CSF2, CTNNB1, DHRS2, DYRK3, KDR, MMP9, PRDX2, TIMP1
immunoglobulin mediated immune response	C2, CD27, PRKCD, SERPING1
inflammatory response to antigenic stimulus	IL20RB
lymphocyte activation	CTNNB1, ITGAM, PRDX2, PRKCD, SLAMF1
lymphocyte differentiation	CTNNB1
lymphocyte mediated immunity	C2, CD27, PRDX1, PRKCD, SERPING1
lymphocyte proliferation	ITGAM, PRDX2, PRKCD
metastasis associated genes	LILRB3, MMP9, PTPN6, TLR6
natural killer cell mediated cytotoxicity	PRDX1
natural killer cell mediated immunity	PRDX1
neutrophil chemotaxis	ITGAM
Cancer Pathways	BAX, CKS1B, CTNNB1, ERBB2, FGF11, ITGA3, MAPK3, MMP9, SMO
response to interferon-gamma	CYP27B1
T cell activation	CTNNB1, ITGAM, PRDX2
T cell differentiation	CTNNB1
T cell differentiation in the thymus	CTNNB1
T cell proliferation	ITGAM, PRDX2
T-cell chronic lymphocytic leukaemia	TCL1A
T-helper 1 type immune response	EBI3, TLR6

TABLE 9 – Cibles potentielles du miR-657 et processus cellulaire associé

# **Conclusion**

Le cancer colorectal est une maladie complexe. Il est le résultat d'un dialogue permanent entre des cellules tumorales et leur environnement. Le système immunitaire joue un grand rôle dans le contrôle de la tumeur. Pour autant, toutes les cellules immunitaires n'ont pas un rôle et un effet identiques. Si une réponse immunitaire de type Th1 ou cytotoxique est de bon pronostic, une réponse de type Th17 est, à l'inverse, défavorable. Il faut également qu'une bonne réponse immunitaire mémoire soit présente. Le type de réponse immunitaire est important mais il faut également que la densité de cellules immunitaires soit suffisante et que ces dernières soient présentes dans les deux régions principales de la tumeur, à savoir le centre et le front de progression. La présence de cellules immunitaires n'est pas suffisante en soi. La réponse immunitaire doit être coordonnée. Ces critères, c'est-à-dire le type, la qualité, la quantité et la coordination, permettent d'identifier les patients ayant un risque de récidiver (figure 28). L'intégration en routine des profils immunologiques aux paramètres cliniques classiques permettrait d'améliorer le diagnostic et la définition des patients à risque.

Il est maintenant nécessaire d'identifier les mécanismes sous-jacents à une bonne réponse immunitaire. Les chimiokines qui attirent des sous-populations lymphocytaires et les microARNs qui modulent la présence d'ARNm, sont autant de pistes. En effet, la chimiokine CX3CL1 attire les lymphocytes T cytotoxiques, tandis que CXCL9 et CXCL10 attirent les lymphocytes T mémoires (figure 28). On peut noter également la corrélation négative de miR-660 avec les fonctions cytotoxiques des cellules, et la corrélation positive des microARNs miR-488 et miR-657 avec la présence de lymphocytes T (figure 28). Leurs mécanismes d'actions restent à préciser.



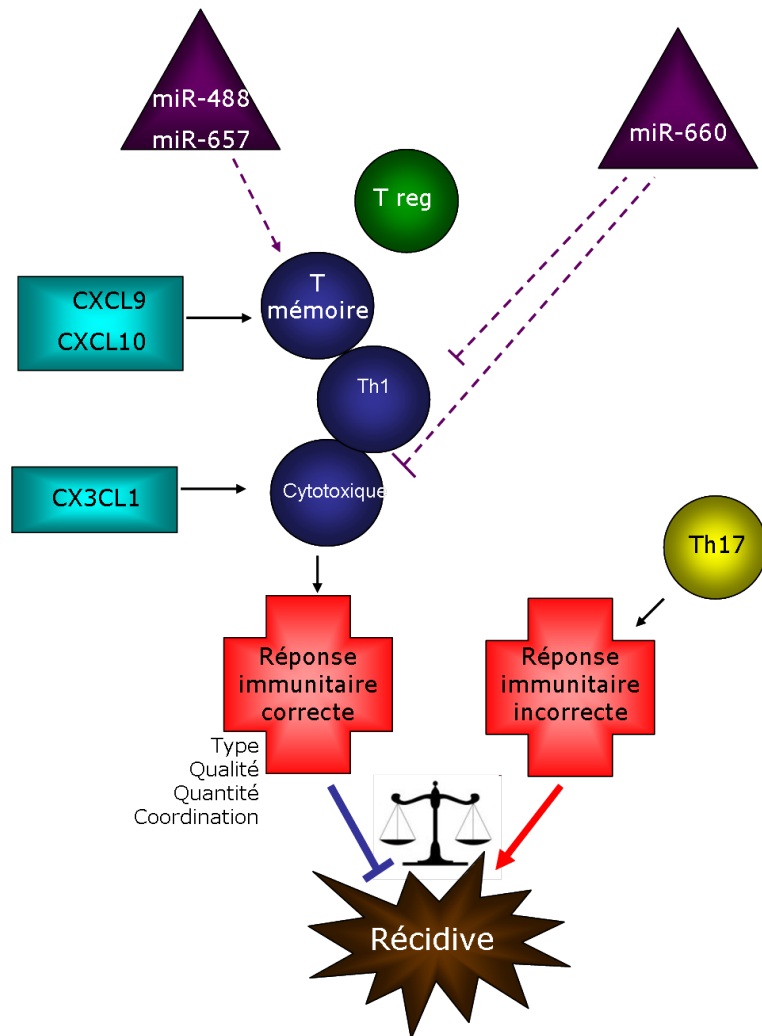


FIGURE 28 – Modèle d'influence du système immunitaire sur la récurrence.

# Bibliographie

- [1] J Weitz, M Koch, J Debus, T Höhler, PR Galle, MW Büchler. Colorectal Cancer. *Lancet*, 365 :153–165, 2005.
- [2] D.C. Chung, A.K. Rustgi. DNA mismatch repair and cancer. *Gastroenterology*, 109 :1687–1699, 1995.
- [3] RJ Davies, R Miller, N Coleman. Colorectal cancer screening : Prospects for molecular stool analysis. *Nature Reviews Cancer*, 5 :199–209, 2005.
- [4] C Dukes. The classification of cancer of the rectum. *J. Pathol.*, 35 :323, 1932.
- [5] E. Mamounas, S. Wieand, et al. Comparative efficacy of adjuvant chemotherapy in patients with Dukes' B versus Dukes' C colon cancer : results from four National surgical Adjuvant Breast and bowel project adjuvant studies. *Journal of clinical oncology*, 5 :1349–1355, 1999.
- [6] B. Vogelstein, ER Fearon, SR Hamilton, SE Kern, AC Preisinger, M. Leppert, Y. Nakamura, R. White, AM Smits, JL Bos. Genetic alterations during colorectal-tumor development. *New England Journal of Medicine*, 319(9) :525–532, 1988.
- [7] H.J. Andreyev, A.R. Norman, D. Cunningham, J.R. Oates, P.A. Clarke. Ras mutations in patients with colorectal cancer : the multicenter "RASCAL" study. *Journal Natl. Cancer Institute*, 90 :675–684, 1998.
- [8] W.S. Samowitz, K. Curtin, D. Schaffer, M. Robertson, M. Leppert, M.L. Slattery. Relationship of K-ras Mutations in Colon Cancers to Tumor Location, Stage, and Survival : A Population-based Study 1. *Cancer Epidemiology Biomarkers & Prevention*, 9(11) :1193–1197, 2000.
- [9] F.A. Monzon, S. Ogino, M.E.H. Hammond, K.C. Halling, K.J Bloom, M.N. Nikiforova. The Role of KRAS Mutation Testing in the Management of Patients With Metastatic Colorectal Cancer. *Gastroenterology*, 133 :1600–1606, 2009.
- [10] A. Walther, E Johnstone, C. Swanton, R Midgley, I. Tomlinson, D. Kerr. Genetic prognosis and predictive markers in colorectal cancer. *Nature Reviews Cancer*, 9 :489–499, 2009.
- [11] S. Popat, R.S. Houlston. A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC and colorectal cancer prognosis. *European Journal of cancer*, 41 :2060–2070, 2005.
- [12] AD Roth, S. Tejpar, P. Yan, et al. Correlation of molecular markers in colon cancer with stage-specific prognosis : results of the translational study on the PETACC3-EORTC40993-SAKK60-00 trial [abstract no. 288]. In *Gastrointestinal Cancers Symposium*, pages 15–17, 2009.
- [13] A. Walther, R. Houlston, I. Tomlinson. Association between chromosomal instability and prognosis in colorectal cancer : a meta-analysis. *Gut*, 57 :941–950, 2008.
- [14] O.J. Finn. Cancer Immunology. *New England Journal of Medicine*, 358 :2704–2715, 2008.
- [15] P. Dalerba, C. Maccalli, C. Casati, C. Castelli, G. Parmiani. Immunology and immunotherapy of colorectal cancer. *Critical Reviews in Oncology and Hematology*, 46(1) :33–57, 2003.
- [16] M.J. Smyth, J. A. Trapani. Lymphocyte-mediated immunosurveillance of epithelial cancers ? *Trends Immunol*, 22 :409–411, 2001.

- [17] M.C. Dieu-Nosjean, M. Antoine, C. Danel, D. Heudes, M. Wislez, V. Poulot, N. Rabbe, L. Laurans, E. Tartour, L. de Chaisemartin, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *Journal of Clinical Oncology*, 26(27) :4410, 2008.
- [18] F. Pages, A. Berger, B. Henglein, B. Piqueras, C. Danel, F. Zinzindohoue, N. Thiounn, P.H. Cugnenc, W.H. Fridman. Modulation of interleukin-18 expression in human colon carcinoma : consequences for tumor immune surveillance. *International Journal of Cancer*, 84(3), 1999.
- [19] F. Pagès, A. Berger, M. Camus, F. Sanchez-Cabo, A. Costes, R. Molitor, B. Mlecnik, A. Kirilovsky, M. Nilsson, D. Damotte, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. *New England Journal of Medicine*, 353(25) :2654–2666, 2005.
- [20] PK Baier, S. Wimmenauer, T. Hirsch, B.U. Specht, S. Kleist, H. Keller, EH Farthmann. Analysis of the T cell receptor variability of tumor-infiltrating lymphocytes in colorectal carcinomas. *Tumor Biol*, 19 :205–212, 1998.
- [21] Y. Naito, K. Saito, K. Shiiba, A. Ohuchi, K. Saigenji, H. Nagura, H. Ohtani. CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Research*, 58(16) :3491–3494, 1998.
- [22] A. Zippelius, P. Batard, V. Rubio-Godoy, G. Bioley, D. Lienard, F. Lejeune, D. Rimoldi, P. Guillaume, N. Meidenbauer, A. Mackensen, et al. Effector function of human tumor-specific CD8 T cells in melanoma lesions : a state of local functional tolerance. *Cancer Research*, 64(8) :2865–2873, 2004.
- [23] N. Auphan-Anezin, C. Mazza, A. Guimezanes, G.A. Barrett-Wilt, F. Montero-Julian, A. Roussel, D.F. Hunt, B. Malissen, A.M. Schmitt-Verhulst. Distinct orientation of the alloreactive monoclonal CD8 T cell activation program by three different peptide/MHC complexes. *European journal of immunology*, 36(7) :1856–1866, 2006.
- [24] T. Boon, P.G. Coulie, B. Van den Eynde. Tumor antigens recognized by T cells. *Immunology Today*, 18(6) :267–268, 1997.
- [25] P. van der Bruggen, Y. Zhang, P. Chaux, V. Stroobant, C. Panichelli, E.S. Schultz, J. Chapiro, B.J. Van den Eynde, F. Brasseur, T. Boon. Tumor-specific shared antigenic peptides recognized by human T cells. *Immunological reviews*, 188(1) :51–64, 2002.
- [26] P. van der Bruggen, C. Traversari, P. Chomez, et al. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, 254 :1643–1647, 1991.
- [27] Y. Zhang, P. Chaux, V. Stroobant, A.M.M. Eggermont, J. Corthals, B. Maillere, K. Thielemans, M. Marchand, T. Boon, P. van der Bruggen. A MAGE-3 peptide presented by HLA-DR1 to CD4+ T cells that were isolated from a melanoma patient vaccinated with a MAGE-3 protein. *The Journal of Immunology*, 171(1) :219, 2003.
- [28] H. Suzuki, D.F. Graziano, J. McKolanis, O.J. Finn. T cell dependant antibody responses against aberrantly expressed cyclin B1 protein in patients with cancer and premalignant disease. *Clinical Cancer Research*, 11 :1521–1526, 2005.
- [29] I. Saeterdal, M.K. Gjertsen, P. Straten, J.A. Eriksen, G. Gaudernack. A TGF betaRII frameshift-mutation-derived CTL epitope recognised by HLA-A2-restricted CD8+ T cells. *Cancer Immunology Immunotherapy*, 50.
- [30] P. Correale, M. Sabatino, MG Cusi, L. Micheli, C. Nencini, D. Pozzessere, R. Petrioli, A. Aquino, L. De Vecchis, M. Turriziani, et al. In vitro generation of cytotoxic T lymphocytes against HLA-A 2. 1-restricted peptides derived from human thymidylate synthase. *Journal of chemotherapy*, 13(5) :519–526, 2001.

- [31] M. Linnebacher, J. Gebert, W. Rudy, S. Woerner, Y.P. Yuan, P. Bork, M. von Knebel Doeberitz. Frameshift peptide-derived T-cell epitopes : a source of novel tumor-specific antigens. *International Journal of Cancer*, 93(1) :6–11, 2001.
- [32] E. Ripberger, M. Linnebacher, Y. Schwitalle, J. Gebert, M.V.K. Doeberitz. Identification of an HLA-A0201-restricted CTL epitope generated by a tumor-specific frameshift mutation in a coding microsatellite of the OGT gene. *Journal of clinical immunology*, 23(5) :415–423, 2003.
- [33] T. Sjoblom, S. Jones, L.D. Wood, D.W. Parsons, J. Lin, T.D. Barber, D. Mandelker, R.J. Leary, J. Ptak, N. Silliman, et al. The consensus coding sequences of human breast and colorectal cancers. *Science*, 314(5797) :268, 2006.
- [34] M. Tsuji, N. Komatsu, S. Kawamoto, K. Suzuki, O. Kanagawa, T. Honjo, S. Hori, S. Fagarasan. Preferential Generation of Follicular B Helper T Cells from Foxp3+ T Cells in Gut Peyer’s Patches, 2009.
- [35] T.R. Mosmann, H. Chenwinski, M.W. Bond, M.A. Gledlin, R.L. Coffman. Two types of murine helper T cell clone : Definition according to profiles of lymphokine activities and secreted proteins. *The Journal of Immunology*, 136 :2348–2357, 1986.
- [36] J. Zhu, W.E. Paul. CD4 T cells : fates, functions and faults. *Blood*, 112 :1557–1569, 2008.
- [37] Y. Chen, V.K. Kuchroo, J. Inobe, D.A. Hafler, Weiner H.L. Regulatory T cell clones induced by oral tolerance : suppression of autoimmune encephalomyelitis. *Science*, 265 :1237–1240, 1994.
- [38] C.L. Langrish, Y. Chen, W.M. Blumenschein, J. Mattson, B. Basham, J.D. Sedgwick, T. McClanahan, R.A. Kastelein, D.J. Cua. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation Abbreviations used : CIA, collagen-induced arthritis ; CNS, central nervous system ; DLN, draining LN ; EAE, experimental autoimmune encephalomyelitis ; MOG, myelin-oligodendrocyte glycoprotein peptide ; PLP, proteolipid protein peptide. *Journal of Experimental Medicine*, 201(2) :233–240, 2005.
- [39] T. Korn, E. Bettelli, M. Oukka, V.K. Kuchroo. IL-17 and Th17 cells. *Annual Revue of Immunology*, 27 :485–517, 2009.
- [40] E.V. Acosta-Rodriguez, G. Napolitani, A. Lanzavecchia, F. Sallusto. Interleukins 1 $\beta$  and 6 but not transforming growth factor  $\beta$  are essential for the differentiation of interleukin 17-producing human T helper cells. *Nature Immunology*, 8 :942–949, 2007.
- [41] N.J. Wilson, K. Boniface, J.R. Chan, B.S. McKenzie, W.M. Blumenschein, J.D. Mattson, B. Basham, K. Smith, T. Chen, F. Morel, et al. Development, cytokine profile and function of human interleukin 17–producing helper T cells. *Nature immunology*, 8(9) :950–957, 2007.
- [42] L. Yang, D.E. Anderson, C. Baecher-Allan, W.D. Hastings, E. Bettelli, M. Oukka, V.K. Kuchroo, D.A. Hafler. IL-21 and TGF- $\beta$  are required for differentiation of human TH17 cells. *Nature*, 454(7202) :350–352, 2008.
- [43] N. Manel, D. Unutmaz, D.R. Littman. The differentiation of human Th17 cells requires TGF $\beta$  and induction of ROR $\gamma$ t. *Nature Immunology*, 9 :641–649, 2008.
- [44] E. Volpe, N. Servant, R. Zollinger, S.I. Bogiatzi, P. Hupé, E. Barillot, V. Soumelis. A critical function for transforming growth factor- $\beta$ , interleukin 23 and proinflammatory cytokines in driving and modulating human TH-17 responses. *Nature immunology*, 9(6) :650–657, 2008.
- [45] N. Fazilleau, L. Mark, L.J. McHeyzer-Williams, M.G. McHeyzer-Williams. Follicular helper T cells : lineage and location. *Immunity*, 30(3) :324–335, 2009.
- [46] J.A. Bluestone, C.R. Mackay, J.J. O’Shea, B. Stockinger. The functional plasticity of T cell subsets. *nature reviews*, 9 :811–816, 2009.

- [47] M. Veldohen, C. Uyttenhove, J. van Snick, H. Helmbj, A. Westendorf, J. Buer, b. Martin, C. Wilhelm, B. Stockinger. Transforming growth factor- $\beta$  'reprograms' the differentiation of T helper 2 cells and promotes an interleukin 9-producing subset. *nature immunology*, 9 :1341–1346, 2008.
- [48] K.S. Voo, Y.H. Wang, F.R. Santori, C. Boggiano, Y.H. Wang, K. Arima, L. Bover, S. Hanabuchi, J. Khalili, E. Marinova, B. Zheng, D.R. Littman, Y.J. Liu. Identification of IL-17-producing FOXP3+ regulatory T cells in humans. *Proceedings of the National Academy of Sciences*, 106 :4793–4789, 2009.
- [49] F. Annunziato, L. Cosmi, F. Liotta, E. Maggi, S. Romagnani. The phenotype of human Th17 cells and their precursors, the cytokines that mediate their differentiation and the role of Th17 cells in inflammation. *International Immunology*, 2008.
- [50] L. Zhou, M.M.W. Chong, D.R. Littman. Plasticity of CD4+ T Cell Lineage Differentiation. *Immunity*, 30 :646–655, 2009.
- [51] A. Glatman-Zaretsky, J.J. Taylor, I.L. King, F.A. Marshall, M. Mohrs, E.J. Pearce. T follicular helper cells differentiate from Th2 cells in response to helminth antigens. *Journal of Experimental Medicine*, 206(5) :991–999, 2009.
- [52] FM Burnet. The concept of immunological surveillance. *Progress in experimental tumor research. Fortschritte der experimentellen Tumorforschung. Progrès de la recherche expérimentale des tumeurs*, 13 :1, 1970.
- [53] L. Zitvogel, A. Tesniere, G. Kroemer. Cancer despite immunosurveillance : immunoselection and immunosubversion. *Nature Reviews Immunology*, 6 :715–727, 2006.
- [54] G.P. Dunn, A.T. Bruce, H. Ikeda, L.J. Old, R.D. Schreiber. Cancer immunoediting : from immunosurveillance to tumor escape. *Nature immunology*, 3(11) :991–998, 2002.
- [55] M.J. Smyth, G.P. Dunn, R.D. Schreiber. Cancer immunosurveillance and immunoediting : the roles of immunity in suppressing tumor development and shaping tumor immunogenicity. *Cancer Immunotherapy*, page 1, 2006.
- [56] A.P. Vicari, C. Caux. Chemokines in cancer. *Cytokine Growth factor revue*, 13 :143–154, 2002.
- [57] P. Matzinger. Tolerance, danger, and the extended family. *Annual Review of Immunology*, 12(1) :991–1045, 1994.
- [58] G.P. Dunn, C.M. Koebel, R.D. Schreiber. Interferons, immunity and cancer immunoediting. *Nature Reviews Immunology*, 6(11) :836–848, 2006.
- [59] E. Di Carlo, G. Forni, P. Lollini, M.P. Colombo, A. Modesti, P. Musiani. The intriguing role of polymorphonuclear neutrophils in antitumor reactions. *Blood*, 97 :339–345, 2001.
- [60] N. Kobayashi. Malignant neoplasms in registered cases of primary immunodeficiency syndrome. *Japanese journal of clinical oncology*, 15 :307, 1985.
- [61] R. Clementi, et al. A proportion of patients with lymphoma may harbor mutations of perforine gene. *Blood*, 105 :4424–4428, 2005.
- [62] M. Carrington, S. Wang, M.P. Martin, X. Gao, M. Schiffman, J. Cheng, R. Herrero, A.C. Rodriguez, R. Kurman, R. Mortel, et al. Hierarchy of resistance to cervical neoplasia mediated by combinations of killer immunoglobulin-like receptor and human leukocyte antigen loci. *Journal of Experimental Medicine*, 201(7) :1069–1075, 2005.
- [63] D. Atkins, A. Breuckmann, G.E. Schmahl, P. Binner, S. Ferrone, F. Krummenauer, S. Storkel, B. Seliger. MHC class I antigen processing pathway defects, ras mutations and disease stage in colorectal carcinoma. *International Journal of Cancer*, 109(2), 2004.
- [64] M.T. Spiotto, P. Yu, D.A. Rowley, M.I. Nishimura, S.C. Meredith, T.F. Gajewski, Y.X. Fu, H. Schreiber. Increasing tumor antigen expression overcomes "ignorance" to solid tumors via crosspresentation by bone marrow-derived stromal cells. *Immunity*, 17(6) :737–747, 2002.

- [65] M. Komatsu, L. Yee, K.L. Carraway. Overexpression of Sialomucin Complex, a Rat Homologue of MUC4, Inhibits Tumor Killing by Lymphokine-activated Killer Cells 1. *Cancer Research*, 59(9) :2229–2236, 1999.
- [66] C. Menetrier-Caux, G. Montmain, MC Dieu, C. Bain, MC Favrot, C. Caux, JY Blay. Inhibition of the differentiation of dendritic cells from CD34+ progenitors by tumor cells : role of interleukin-6 and macrophage colony-stimulating factor. *Blood*, 92(12) :4778, 1998.
- [67] D.I. Gabrilovich, H.L. Chen, K.R. Girgis, H.T. Cunningham, G.M. Meny, S. Nadaf, D. Kavanaugh, D.P. Carbone. Production of vascular endothelial growth factor by human tumors inhibits the functional maturation of dendritic cells. *Nature medicine*, 2(10) :1096–1103, 1996.
- [68] T. Joki, O. Heese, D.C. Nikas, L. Bello, J. Zhang, S.K. Kraeft, N.T. Seyfried, T. Abe, L.A.N.B.O. Chen, R.S. Carroll, et al. Expression of cyclooxygenase 2(COX-2) in human glioma and in vitro inhibition by a specific COX-2 inhibitor, NS-398. *Cancer research*, 60(17) :4926–4931, 2000.
- [69] C.C. Sombroek, A.G.M. Stam, A.J. Masterson, S.M. Lougheed, M.J.A.G. Schakel, C.J.L.M. Meijer, H.M. Pinedo, A.J.M. Van den Eertwegh, R.J. Scheper, T.D. Gruijl. Prostanoids play a major role in the primary tumor-induced inhibition of dendritic cell differentiation. *The Journal of Immunology*, 168(9) :4333, 2002.
- [70] H. Harizi, M. Juzan, C. Grosset, M. Rashedi, N. Gualde. Dendritic cells issued in vitro from bone marrow produce PGE2 that contributes to the immunomodulation induced by antigen-presenting cells. *Cellular Immunology*, 209(1) :19–28, 2001.
- [71] G.V. Shurin, M.R. Shurin, S. Bykovskaia, J. Shogan, M.T. Lotze, E.M. Barksdale. Neuroblastoma-derived Gangliosides Inhibit Dendritic Cell Generation and Function 1. *Cancer Research*, 61(1) :363–369, 2001.
- [72] C. Uyttenhove, L. Pilotte, I. Théate, V. Stroobant, D. Colau, N. Parmentier, T. Boon, B.J. Van den Eynde. Evidence for a tumoral immune resistance mechanism based on tryptophan degradation by indoleamine 2, 3-dioxygenase. *Nature medicine*, 9(10) :1269–1274, 2003.
- [73] H. Dong, S.E. Strome, D.R. Salomao, H. Tamura, F. Hirano, D.B. Flies, P.C. Roche, J. Lu, G. Zhu, K. Tamada, et al. Tumor-associated B7-H1 promotes T-cell apoptosis : a potential mechanism of immune evasion. *Nature medicine*, 8(8) :793–800, 2002.
- [74] M. Nakashima, K. Sonoda, T. Watanabe. Inhibition of cell growth and induction of apoptotic cell death by the human tumor-associated antigen RCAS1. *Nature medicine*, 5(8) :938–942, 1999.
- [75] JP Medema, J. De Jong, LTC Peltenburg, EME Verdegaal, A. Gorter, SA Bres, K. Franken, M. Hahne, JP Albar, CJM Melief, et al. Blockade of the granzyme B/perforin pathway through overexpression of the serine protease inhibitor PI-9/SPI-6 constitutes a mechanism for immune escape by tumors. *Proceedings of the National Academy of Sciences*, 98(20) :11515, 2001.
- [76] S. Chouaib, F. Meslin, J. Thiery, F. Mami-Chouaib. Tumor resistance to specific lysis : A major hurdle for successful immunotherapy of cancer. *Clinical Immunology*, 130(1) :34–40, 2009.
- [77] D. Daniel, N. Meyer-Morse, E.K. Bergsland, K. Dehne, L.M. Coussens, D. Hanahan. Immune enhancement of skin carcinogenesis by CD4+ T cells. *Journal of Experimental Medicine*, 197(8) :1017–1028, 2003.
- [78] C. Conticello, F. Pedini, A. Zeuner, M. Patti, M. Zerilli, G. Stassi, A. Messina, C. Peschle, R. De Maria. IL-4 Protects Tumor Cells from Anti-CD95 and Chemotherapeutic Agents via Up-Regulation of Antiapoptotic Proteins 1. *The Journal of Immunology*, 172(9) :5467–5477, 2004.
- [79] J.J. Oppenheim, H.F. Dong, P. Plotz, R.R. Caspi, M. Dykstra, S. Pierce, R. Martin, C. Carlos, O. Finn, O. Koul, et al. Autoantigens act as tissue-specific chemoattractants. *Journal of leukocyte biology*, 77(6) :854–861, 2005.

- [80] T.J. Curiel, P. Cheng, P. Mottram, X. Alvarez, L. Moons, M. Evdemon-Hogan, S. Wei, L. Zou, I. Kryczek, G. Hoyle, et al. Dendritic cell subsets differentially regulate angiogenesis in human ovarian cancer. *Cancer Research*, 64(16) :5535–5538, 2004.
- [81] Y. Itoh, T. Joh, S. Tanida, M. Sasaki, H. Kataoka, K. Itoh, T. Oshima, N. Ogasawara, S. Togawa, T. Wada, et al. IL-8 promotes cell proliferation and migration through metalloproteinase-cleavage proHB-EGF in human colon carcinoma cells. *Cytokine*, 29(6) :275–282, 2005.
- [82] RC. Lee, R.L. Feinbaum, V. Ambros. The *C. elegans* heterochronic gene *lin-14* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75 :843, 1993.
- [83] BJ. Reinhart, FJ. Slack, M. Basson, Pasquinelli AE., JC. Bettinger, AE. Rougvie, et al. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 403 :901–906, 2000.
- [84] N.C. Lau, L.P. Lim, E.G. Weinstein, D.P. Bartel. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*, 294 :858–862, 2001.
- [85] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschli. Identification of novel genes encoding for small expressed RNAs. *Science*, 294 :853–858, 2001.
- [86] MT. McManus. MicroRNAs and cancer. *Seminars in cancer biology*, 13 :253–258, 2003.
- [87] AE. Pasquinelli, BJ. Reinhart, FJ. Slack, MQ. Martindale, MJ. Kuroda, B. Maller, et al. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature*, 408 :86–89, 2000.
- [88] A. Ventura, T. Jacks. MicroRNAs and cancer : short RNAs go a long way. *Cell*, 136(4) :586–591, 2009.
- [89] D.P. Bartel. MicroRNAs genomics, biogenesis, mechanism, and function. *Cell*, 116(2) :281–297, 2004.
- [90] F. Ozsolak, L.L. Poling, Z. Wang, H. Liu, X.S. Liu, R.G. Roeder, X. Zhang, J.S. Song, D.E. Fisher. Chromatin structure analyses identify miRNA promoters. *Genes & Development*, 22(22) :3172, 2008.
- [91] C.P. Petersen, M.E. Bordeleau, J. Pelletier, P.A. Sharp. Short RNAs repress translation after initiation in mammalian cells. *Molecular cell*, 21(4) :533–542, 2006.
- [92] J.M. Thomson, M. Newman, J.S. Parker, E.M. Morin-Kensicki, T. Wright, S.M. Hammond. Extensive post-transcriptional regulation of microRNAs and its implication for cancer. *Genes Development*, 20 :2202–2207, 2006.
- [93] S. Yekta, IH. Shih, D.P. Bartel. microRNA-directed cleavage of *HOXB8* mRNA. *Science*, 304 :594–596, 2004.
- [94] S. Vasudevan, Y. Tong, J.A. Steitz. Switching from repression to activation : microRNAs can up-regulate translation. *Science*, 318(5858) :1931, 2007.
- [95] D.H. Kim, P. Sætrom, O. Snøve, J.J. Rossi. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proceedings of the National Academy of Sciences*, 105(42) :16230, 2008.
- [96] F. Robert, L.C. Li, D. Pookot, E.J. Noonan, R. Dahiya. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences*, 105(5), 2008.
- [97] K.A. O'Donnell, E.A. Wentzel, K.I. Zeller, C.V. Dang, J.T. Mendell. c-Myc-regulated microRNAs modulate *E2F1* expression. *Nature*, 435 :839–843, 2005.
- [98] Y. Kawahara, B. Zinshteyn, P. Sethupathy, H. Iizasa, A.G. Hatzigeorgiou, K. Nishikura. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science*, 315(5815) :1137, 2007.
- [99] S.W. Knight, B.L. Bass. The role of RNA editing by ADARs in RNAi. *Molecular Cell*, 10 :809–817, 2002.
- [100] W. Yang, T.P. Chendrimada, Q. Wang, M. Higuchi, P.H. Seeburg, R. Shiekhattar, K. Nishikura. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature structural molecular biology*, 13 :13–21, 2006.

- [101] S.I. Ashraf, A.L. McLoon, S.M. Sclasic, S. Kunes. Synaptic protein synthesis associated with memory is regulated by RISC pathway in *Drosophila*. *Cell*, 124 :191–205, 2006.
- [102] M. Kedde, M.J. Strasser, B. Boldajipour, J.A.F.O. Vrieling, K. Slanchev, C. le Sage, R. Nagel, P.M. Voorhoeve, J. van Duijse, U.A. Ørom, et al. RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell*, 131(7) :1273–1286, 2007.
- [103] S.N. Bhattacharyya, R. Habermacher, U. Martine, E.I. Closs, W. Filipowicz. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell*, 125(6) :1111–1124, 2006.
- [104] F. Fazi, A. Rosa, V. Gelmetti, M.L. DeMarchis, C. Nervis, I. Bozzoni. A minicircuitry comprised of microRNA-223 and transcription factor NFI-A and C/EBP $\alpha$  regulates human granulopoiesis. *Cell*, 123 :819–831, 2005.
- [105] J.B. Johnnidis, M.H. Harris, R.T. Wheeler, S. Stehling-Sun, M.H. Lam, O. Kirak, T.R. Brummelkamp, M.D. Fleming, F.D. Camargo. Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature*, 451 :1125–1129, 2008.
- [106] D. Baltimore, M.P. Boldin, R.M. O’Connell, D.S. Rao, K.D. Taganov. MicroRNAs : new regulators of immune cell development and function. *Nature Immunology*, 9(8) :839–845, 2008.
- [107] Y. Bi, G. Liu, R. Yang. MicroRNAs : novel regulators during the immune response. *Journal of Cellular Physiology*, 218(3), 2009.
- [108] E. Sonkoly, M. Ståhle, A. Pivarcsi. MicroRNAs and immunity : novel players in the regulation of normal immune function and inflammation. In *Seminars in cancer biology*, volume 18, pages 131–140. Elsevier, 2008.
- [109] K. Basso, P. Sumazin, P. Morozov, C. Schneider, R.L. Maute, Y. Kitagawa, J. Mandelbaum, J. Haddad, C.Z. Chen, A. Califano, et al. Identification of the Human Mature B Cell miRNome. *Immunity*, 2009.
- [110] T.H. Thai, D.P. Calado, S. Casola, K.M. Ansel, C. Xiao, Y. Xue, A. Murphy, D. Friendewey, D. Valenzuela, J.L. Kutok, et al. Regulation of the germinal center response by microRNA-155. *Science*, 316(5824) :604, 2007.
- [111] E. Vigorito, K.L. Perks, C. Abreu-Goodger, S. Bunting, Z. Xiang, S. Kohlhaas, P.P. Das, E.A. Miska, A. Rodriguez, A. Bradley, et al. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity*, 27 :847–859, 2007.
- [112] R. Schickel, B. Boyerinas, S.M. Park, M.E. Peter. MicroRNAs : key players in the immune system, differentiation, tumorigenesis and cell death. *Oncogene*, 27(45) :5959–5974, 2008.
- [113] J.B. COWLAND, C. HOTHER, K. GRONBAEK. MicroRNAs and cancer. *Apmis*, 115(10) :1090–1106, 2007.
- [114] C. Kanellopoulou, S. Monticelli. A role for microRNAs in the development of the immune system and in the pathogenesis of cancer. In *Seminars in Cancer Biology*, volume 18, pages 79–88. Elsevier, 2008.
- [115] Z. Zhang, Z. Li, C. Gao, J. Chen, W. Liu, S. Xiao, H. Lu. miR-21 plays a pivotal role in gastric cancer pathogenesis and progression. *Laboratory Investigation*, 88 :1358–1366, 2009.
- [116] Y. Ladeiro, G. Couchy, C. Balabaud, P. Bioulac-Sage, L. Pelletier, S. Rebouissou, J. Zucman-Rossi. MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. Potential conflict of interest : Nothing to report. *Hepatology*, 47(6) :1955–1963, 2008.
- [117] J.T. Mendell. miRiad roles for the miR-17-92 in development and disease. *Cell*, 133 :217–222, 2008.
- [118] Y. Sylvestre, V. De Guire, E. Querido, U.K. Mukhopadhyay, V. Bourdeau, F. Major, G. Ferbeyre, P. Chartrand. An E2F/miR-20a autoregulatory feedback loop. *Journal of Biological Chemistry*, 282 :2135–2143, 2007.



- [119] K. Woods, J.M. Thomson, S.M. Hammond. Direct regulation of an oncogenic microRNA cluster by E2F transcription factor. *Journal of Biological chemistry*, 282 :2130–2134, 2007.
- [120] I. Ivanovska, A.S. Ball, R.L. Diaz, M. Magnus, J.F. and Kibukawa, J.M. Schelter, S.V. Kobayashi, L. Lim, J. Burchard, A.L. Jackson, et al. MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression. *Molecular and cellular biology*, 28 :2167–2174, 2008.
- [121] D. Baltimore, M.P. Boldin, R.M. O’Connell, D.S. Rao, K.D. Taganov. Lymphoproliferative disease and autoimmunity in mice with elevated miR-17-92 expression in lymphocytes. *Nature Immunology*, 9 :405–414, 2008.
- [122] A. Hossain, M.T. Kuo, G.F. Saunders. MiR-17-5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA. *Molecular and cellular biology*, 26 :8191–8201, 2006.
- [123] M. Dews, A. Homayouni, D. Yu, D. Murphy, C. Sevignani, E. Wentzel, E.E. Furth, W.M. Lee, G.H. Enders, J.T. Mendell, A. Thomas-Tikhonenko. Augmentation of tumor angiogenesis by Myc-activated microRNA cluster. *Nature Genetics*, 38 :1060–1065, 2006.
- [124] A. Bonauer, G. Carmona, M. Iwasaki, M. Mione, M. Koyanagi, A. Fischer, J. Burchfield, H. Fox, C. Doebele, K. Ohtani, et al. MicroRNA-92a Controls Angiogenesis and Functional Recovery of Ischemic Tissues in Mice. *Science*, 324(5935) :1710, 2009.
- [125] M.I. Aslam, K. Taylor, J.H. Pringle, J.S. Jameson. MicroRNAs are biomarkers of colorectal cancer. *British Journal of Surgery*, 96 :702–710, 2009.
- [126] IA Asangani, SAK Rasheed, DA Nikolova, JH Leupold, NH Colburn, S. Post, H. Allgayer. MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene*, 27(15) :2128–2136, 2007.
- [127] R. Diaz, J. Silva, J.M. Garcia, Y. Lorenzo, V. Garcia, C. Pena, R. Rodriguez, C. Munoz, F. Garcia, F. Bonilla, et al. Deregulated expression of miR-106a predicts survival in human colon cancer patients. *Genes, Chromosomes and Cancer*, 47(9), 2008.
- [128] D.H. Yasui, T. Genetta, T. Kadesch, T.M. Williams, S.L. Swain, L.V. Tsui, B.T. Huber. Transcriptional Repression of the IL-2 Gene in Th Cells by ZEB 1. *The Journal of Immunology*, 160(9) :4433–4440, 1998.
- [129] H. Xu, I.Y. Cheung, H.F. Guo, N.K.V. Cheung. microRNA miR-29 modulates expression of immunoinhibitory molecule B7-H3 : potential implications for immune based therapy of human solid tumors. *Cancer research*, 69(15) :6275–6281, 2009.
- [130] N. Bandi, S. Zbinden, M. Gugger, M. Arnold, V. Kocher, L. Hasan, A. Kappeler, T. Brunner, E. Vassella. miR-15a and miR-16 Are Implicated in Cell Cycle Regulation in a Rb-Dependent Manner and Are Frequently Deleted or Down-regulated in Non-Small Cell Lung Cancer. *Cancer Research*, 69(13) :5553, 2009.
- [131] S. Valastyan, F. Reinhardt, N. Benaich, D. Calogrias, A.M. Szász, Z.C. Wang, J.E. Brock, A.L. Richardson, R.A. Weinberg. A Pleiotropically Acting MicroRNA, miR-31, Inhibits Breast Cancer Metastasis. *Cell*, 137(6) :1032–1046, 2009.
- [132] X. Chen, X. Guo, H. Zhang, Y. Xiang, J. Chen, Y. Yin, X. Cai, K. Wang, G. Wang, Y. Ba, et al. Role of miR-143 targeting KRAS in colorectal tumorigenesis. *Oncogene*, 28(10) :1385–1392, 2009.
- [133] S.R. Viswanathan, G.Q. Daley, R.I. Gregory. Selective blockade of microRNA processing by Lin-28. *Science*, 320 :97–100, 2008.
- [134] S.M. Johnson, H. Grosshans, J. Shingara, M. Byrom, R. Jarvis, A. Cheng, E. Labourier, K.L. Reinert, D. Brown, F.J. Slack. RAS is regulated by the let-7 microRNA family. *Cell*, 120(5) :635–647, 2005.

- [135] Y.S. Lee, A. Dutta. The tumor repressor microRNA let-7 represses theHMGA2 oncogene. *genes development*, 21 :1025–1030, 2007.
- [136] V.B. Sampson, N.H. Rong, J. Han, Q. Yang, V. Aris, P. Soteropoulos, N.J. Petrelli, S.P. Dunn, L.J. Krueger. MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer research*, 67(20) :9762, 2007.
- [137] F. Meng, R. Henson, H. Wehbe-Janek, H. Smith, Y. Ueno, T. Patel. The microRNA let-7a modulates interleukin-6-dependant STAT3 survival signaling in malignant human cholangiocytes. *Journal of Biological chemistry*, 282 :8256–8264, 2007.
- [138] T. Schepeler, J.T. Reinert, M.S. Ostensfeld, L.L. Christensen, A.N. Silahiroglu, L. Dyrskjot, C. Wiuf, F.J. Sorensen, M. Kruhoffer, S. Laurberg, et al. Diagnostic and prognostic microRNAs in stage II colon cancer. *Cancer Research*, 68(15) :6416, 2008.
- [139] B. Shi, L. Sepp-Lorenzino, M. Prisco, P. Linsley, T. de Angelis, R. Baserga. MicroRNA 145 targets the insulin receptor substrate 1 and inhibits the growth of colon cancer. *Journal of Biological chemistry*, 282 :32582–32590, 2007.
- [140] J. Galon, A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pages, M. Tosolini, M. Camus, A. Berger, P. Wind, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*, 313(5795) :1960–1964, 2006.
- [141] M. Camus, M. Tosolini, B. Mlecnik, F. Pages, A. Kirilovsky, A. Berger, A. Costes, G. Bindea, P. Charoentong, P. Bruneval, et al. Coordination of Intratumoral Immune Reaction and Human Colorectal Cancer Recurrence. *Cancer research*, 69(6) :2685, 2009.
- [142] C. Loddenkemper, M. Schernus, M. Noutsias, H. Stein, E. Thiel, D. Nagorsen. In situ analysis of FOXP3+ regulatory T cells in human colorectal cancer. *Journal of translational medicine*, 4(52), 2006.
- [143] P. Salama, M. Phillips, F. Grieu, M. Morris, N. Zeps, D. Joseph, C. Platell, B. Iacopetta. Tumor-infiltrating FOXP3+ T regulatory cells show strong prognostic significance in colorectal cancer. *Journal of Clinical Oncology*, 27(2) :186, 2009.
- [144] H.T. Khong, N.P. Restifo. natural selection of tumor variants in the generation of "tumor escape" phenotype. *Nature Immunology*, 3(11) :999–1005, 2002.
- [145] L.M. Coussens, Z. Werb. Inflammation and cancer. *Nature*, (420) :860–867, 2002.
- [146] S. Pizzimenti, M. Ferracin, S. Sabbioni, C. Toaldo, P. Pettazoni, M.U. Dianzani, M. Negrini, G. Barrera. MicroRNA expression changes during human leukemic HL-60 cell differentiation induced by 4-hydroxynonenal, a product of lipid peroxidation. *Free Radical Biology and Medicine*, 46(2) :282–288, 2009.
- [147] K. Lv, Y. Guo, Y. Zhang, K. Wang, Y. Jia, S. Sun. Allele-specific targeting of hsa-miR-657 to human IGF2R creates a potential mechanism underlying the association of ACAA-insertion/deletion polymorphism with type 2 diabetes. *Biochemical and Biophysical Research Communications*, 374(1) :101–105, 2008.

# **Annexes**

## **Annexe 1**

# **"ClueGO : a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks"**

Gabriela Bindea, Bernhard Mlecnik, Hubert Hackl, Pornpimol Charoentong, Marie Tosolini, Amos Kirilovsky, Wolf-Herman Fridman, Franck Pagès, Zlatko Trajanoski and Jérôme Galon

Bioinformatics, 2009, volume 23, pages 1091-1093

*Systems biology*

## ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks

Gabriela Bindea<sup>1,2,3,4,6</sup>, Bernhard Mlecnik<sup>1,2,3,6</sup>, Hubert Hackl<sup>4</sup>, Pornpimol Charoentong<sup>4</sup>, Marie Tosolini<sup>1,2,3</sup>, Amos Kirilovsky<sup>1,2,3</sup>, Wolf-Herman Fridman<sup>1,2,3,5</sup>, Franck Pagès<sup>1,2,3,5</sup>, Zlatko Trajanoski<sup>4</sup> and Jérôme Galon<sup>1,2,3,5</sup> \*

<sup>1</sup>INSERM, AVENIR Team, Integrative Cancer Immunology, U872, 75006 Paris, France; <sup>2</sup>Université Paris Descartes, Paris, France; <sup>3</sup>Université Pierre et Marie Curie Paris 6, Cordeliers Research Center, Paris, France; <sup>4</sup>Institute for Genomics and Bioinformatics, Graz University of Technology, Graz, Austria; <sup>5</sup>Assistance Publique-Hôpitaux de Paris, HEGP, Paris, France; <sup>6</sup>Equal contribution.

Received on November 13, 2008; revised on January 24, 2009;

Associate Editor: Trey Ideker

### ABSTRACT

**Summary:** We have developed ClueGO, an easy to use Cytoscape plug-in that strongly improves biological interpretation of large lists of genes. ClueGO integrates Gene Ontology (GO) terms as well as KEGG/BioCarta pathways and creates a functionally organized GO/pathway term network. It can analyze one or compare two lists of genes and comprehensively visualizes functionally grouped terms. A one-click update option allows ClueGO to automatically download the most recent GO/KEGG release at any time. ClueGO provides an intuitive representation of the analysis results and can be optionally used in conjunction with the Golorize plug-in.

**Availability:** <http://www.ici.upmc.fr/cluego/cluegoDownload.shtml>

**Contact:** [jerome.galon@crc.jussieu.fr](mailto:jerome.galon@crc.jussieu.fr)

**Supplementary information:**

<http://www.ici.upmc.fr/cluego/ClueGODocumentation.pdf>

GO terms and reconstruct the hierarchical ontology tree, whereas ClueGO uses kappa statistics to link the terms in the network. Compared to the approach of (Ramos *et al.*, 2008) which creates an *in silico* annotation network based on pathways and protein interaction data and maps the gene list of interest afterwards, ClueGO generates a dynamical network structure by already initially considering the gene lists of interest. ClueGO integrates GO terms as well as KEGG/BioCarta pathways and creates a functionally organized GO/pathway term network. A variety of flexible restriction criteria allow for visualizations in different levels of specificity. In addition, ClueGO can compare clusters of genes and visualizes their functional differences. ClueGO takes advantage of Cytoscape's versatile visualization framework and can be used in conjunction with the Golorize plug-in (Garcia *et al.*, 2007).

## 1 INTRODUCTION

Since the number of genes that can be analyzed by high-throughput experiments by far exceeded what can be interpreted by a single person, different attempts have been initiated in order to capture biological information and systematically organize the wealth of data. For example Gene Ontology (GO) (Ashburner *et al.*, 2000) annotates genes to biological/cellular/molecular terms in a hierarchically structured way, whereas Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa *et al.*, 2002) and BioCarta assigns genes to functional pathways. Several functional enrichment analysis tools (e.g. Zeeberg *et al.*, 2003; Boyle *et al.*, 2004; Maere *et al.*, 2005; Huang *et al.*, 2007; Ramos *et al.*, 2008)) and algorithms (e.g. (Li *et al.*, 2008)) were developed to enhance data interpretation.

As most of these tools mainly present their results as long lists or complex hierarchical trees, we aimed to develop ClueGO a Cytoscape (Shannon *et al.*, 2003) plug-in to facilitate the biological interpretation and to visualize functionally grouped terms in the form of networks and charts. Other tools like BiNGO (Maere *et al.*, 2005) or PIPE (Ramos *et al.*, 2008) assess overrepresented

## 2 METHODS AND IMPLEMENTATION

ClueGO has two major features: It can be either used for the visualization of terms corresponding to a list of genes, or the comparison of functional annotations of two clusters.

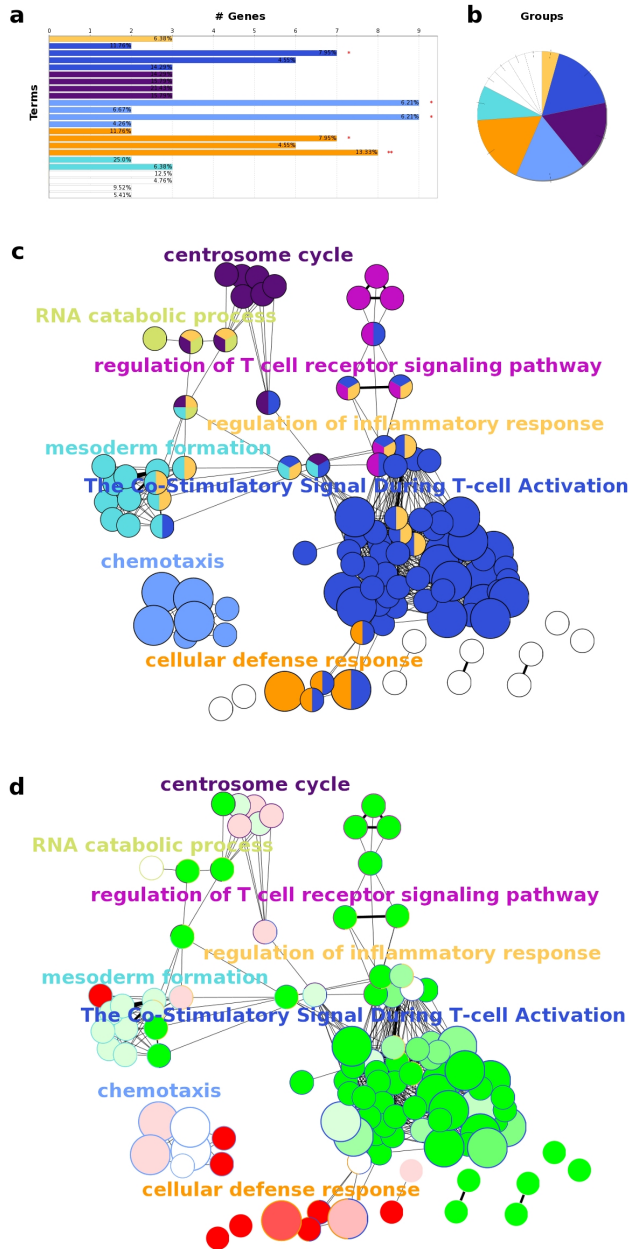
### Data import

Gene identifier sets can be directly uploaded in simple text format or interactively derived from gene network graphs visualized in Cytoscape. ClueGO supports several gene identifiers and organisms by default and is easy extendable for additional ones in a plug-in like manner (see supplementary information).

### Annotation sources

To allow a fast analysis, ClueGO uses precompiled annotation files including GO, KEGG and BioCarta for a wide range of organisms. A one-click update feature automatically downloads the latest ontology and annotation sources and creates new precompiled files that are added to the existing ones. This ensures an up-to-date functional analysis. Additionally ClueGO can easily integrate new annotation sources in a plug-in like way (see supplementary information).

\*to whom correspondence should be addressed



**Fig. 1.** ClueGO example analysis of up and down regulated NK cell genes in peripheral blood from healthy human donors. a) GO/pathway terms specific for up regulated genes. The bars represent the number of genes associated with the terms. The percentage of genes per term is shown as bar label. b) Overview chart with functional groups including specific terms for up regulated genes. c) Functionally grouped network with terms as nodes linked based on their kappa score level ( $\geq 0.3$ ), where only the label of the most significant term per group is shown. The node size represents the term enrichment significance. Functionally related groups partially overlap. Not grouped terms are shown in white. d) The distribution of two clusters visualized on network c). Terms with up/down regulated genes are shown in red/green, respectively. The color gradient shows the gene proportion of each cluster associated with the term. Equal proportions of the two clusters are represented in white.

### Enrichment tests

ClueGO offers the possibility to calculate enrichment/depletion tests for terms and groups as left-sided (Enrichment), right-sided (Depletion) or two-sided (Enrichment/Depletion) tests based on the hypergeometric distribution. Furthermore it provides options to calculate mid- $P$ -values and doubling for two-sided tests to deal with discreteness and conservatism effects as suggested by (Rivals *et al.*, 2007). To correct the  $P$ -values for multiple testing several standard correction methods are proposed (Bonferroni, Bonferroni step-down and Benjamini-Hochberg).

### Network generation and visualization

To create the annotations network ClueGO provides predefined functional analysis settings ranging from general to very specific ones. Furthermore, the user can adjust the analysis parameters to focus on terms e.g. in certain GO level intervals, with particular evidence codes or with a certain number and percentage of associated genes. An optional redundancy reduction feature (Fusion) assesses GO terms in a parent - child relation sharing similar associated genes and preserves the more representative parent or child term. The relationship between the selected terms is defined based on their shared genes in a similar way as described by (Huang *et al.*, 2007). ClueGO creates first a binary gene-term matrix with the selected terms and their associated genes. Based on this matrix, a term-term similarity matrix is calculated using chance corrected kappa statistics to determine the association strength between the terms. Since the term-term matrix is of categorical origin, kappa statistic was found to be the most suitable method. Finally the created network represents the terms as nodes which are linked based on a predefined kappa score level. The kappa score level threshold can initially be adjusted on a positive scale from 0 to 1 to restrict the network connectivity in a customized way. The size of the nodes reflects the enrichment significance of the terms. The network is automatically laid out using the Organic layout algorithm supported by Cytoscape. The functional groups are created by iterative merging of initially defined groups based on the predefined kappa score threshold. The final groups are fixed or randomly colored and overlaid with the network. Functional groups represented by their most significant (leading) term are visualized in the network providing an insightful view of their interrelations. Also other ways of selecting the group leading term e.g. based on the number or percentage of genes per term are provided. As an alternative to the kappa score grouping the GO hierarchy using parent-child relationships can be used to create functional groups.

When comparing two gene clusters, another original feature of ClueGO allows to switch the visualization of the groups on the network to the cluster distribution over the terms. Besides the network, ClueGO provides overview charts showing the groups and their leading term as well as detailed term histograms for both, cluster specific and common terms.

Like BiNGO, ClueGO can be used in conjunction with Golorize for functional analysis of a Cytoscape gene network. The created networks, charts and analysis results can be saved as project in a specified folder and used for further analysis.

## 3 CASE STUDY

To demonstrate how ClueGO assesses and compares biological functions for clusters of genes we selected up and down regulated

natural killer (NK) cell genes in healthy donors from an expression profile of human peripheral blood lymphocytes (GSE6887, Gene Expression Omnibus). For up regulated NK genes ClueGO revealed specific terms like "Natural killer cell mediated cytotoxicity" in the group "Cellular defense response". Down regulated in NK cells compared with the reference (a pool of all immune cell types) were genes involved in the innate immune response (Macrophages) but also in the adaptive immune response (T and B cell). The common functionality refers to characteristics of leukocytes (chemotaxis), besides other terms involved in cell division and metabolism (Fig.1).

#### 4 SUMMARY

ClueGO is a user friendly Cytoscape plug-in to analyze interrelations of terms and functional groups in biological networks. A variety of flexible adjustments allow for a profound exploration of gene clusters in annotation networks. Our tool is easily extendable to new organisms and identifier types as well as new annotation sources which can be included in a transparent, plug-in like manner. Furthermore, the one-click update feature of ClueGO ensures an up-to-date analysis at any time.

#### ACKNOWLEDGMENTS

We thank A Van Cortenbosch for the name of the tool. This work was supported by INSERM, Ville de Paris, INCa, the Austrian

Ministry for Science and Research, Project GEN-AU, BINII and the European 7FP, Grant Agreement 202230 (GENINCA).

#### REFERENCES

- Ashburner,M. *et al.* (2000). Gene ontology: tool for the unification of biology The Gene Ontology Consortium. *Nat Genet*, **25**, 25–29.
- Boyle,E. I. *et al.* (2004). GO::TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics*, **20**, 3710–3715.
- Garcia,O. *et al.* (2007). Golorize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. *Bioinformatics*, **23**, 394–396.
- Huang,d. W. *et al.* (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*, **8**, R183–R183.
- Kanehisa,M. *et al.* (2002). The KEGG databases at GenomeNet. *Nucleic Acids Res*, **30**, 42–46.
- Li,Y. *et al.* (2008). A global pathway crosstalk network. *Bioinformatics*, **24**, 1442–1447.
- Maere,S. *et al.* (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics*, **21**, 3448–3449.
- Ramos,H. *et al.* (2008). The protein information and property explorer: an easy-to-use, rich-client web application for the management and functional analysis of proteomic data. *Bioinformatics*, **24**, 2110–2111.
- Rivals,I. *et al.* (2007). Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics*, **23**, 401–407.
- Shannon,P. *et al.* (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, **13**, 2498–2504.
- Zeeberg,B. R. *et al.* (2003). GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol*, **4**, R28–R28.

## **Annexe 2**

# **"The in situ Cytotoxic and Memory T Cells Predict Outcome in Early-Stage Colorectal Cancer Patients"**

Franck Pagès\*, Amos Kirilovsky\*, Bernhard Mlecnik, Martin Asslaber, Marie Tosolini, Gabriela Bindea, Christine Lagorce, Philippe Wind, Patrick Bruneval, Kurt Zatloukal, Zlatko Trajanoski, Anne Berger, Wolf-Herman Fridman, and Jérôme Galon

Journal of clinical oncology, 2009, volume 27, numéro 35, page 5944-5951



## In Situ Cytotoxic and Memory T Cells Predict Outcome in Patients With Early-Stage Colorectal Cancer

Franck Pagès, Amos Kirilovsky, Bernhard Mlecnik, Martin Asslaber, Marie Tosolini, Gabriela Bindea, Christine Lagorce, Philippe Wind, Florence Marliot, Patrick Bruneval, Kurt Zatloukal, Zlatko Trajanoski, Anne Berger, Wolf-Herman Fridman, and Jérôme Galon

From the AVENIR L'Institut National de la Santé et de la Recherche Médicale (INSERM) Integrative Cancer Immunology Team, INSERM U872; Université Paris-Descartes; Cordeliers Research Center; Université Pierre et Marie Curie Paris 6; Departments of Immunology, Pathology, and General and Digestive Surgery of the Georges Pompidou European Hospital, Assistance Publique Hôpitaux de Paris, Paris; Departments of General and Digestive Surgery, and Pathology, Avicenne Hospital, Bobigny, France; Institute for Genomics and Bioinformatics, Graz University of Technology; and Institute of Pathology, Medical University of Graz, Graz, Austria.

Submitted August 13, 2008; accepted July 17, 2009; published online ahead of print at [www.jco.org](http://www.jco.org) on October 26, 2009.

Supported by L'Institut National de la Santé et de la Recherche Médicale, the National Cancer Institute, Association pour la Recherche sur le Cancer, the Cancéropole Ile de France, Ville de Paris, Immuncan, Austrian Academy of Science, the Austrian Federal Ministry of Science and Research, Bioinformatics Integration Network Austrian Genome Programme (GEN-AU), the GEN-AU, and by the European Commission (FP7, Geninca Consortium, Grant No. 202230).

F.P. and A.K. contributed equally to this work.

Terms in blue are defined in the glossary, found at the end of this article and online at [www.jco.org](http://www.jco.org).

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Jérôme Galon, PhD, INSERM AVENIR, U872, 15 rue de l'Ecole de Médecine, Centre de Recherche des Cordeliers, 75006 Paris, France; e-mail: [jerome.galon@crc.jussieu.fr](mailto:jerome.galon@crc.jussieu.fr).

© 2009 by American Society of Clinical Oncology

0732-183X/09/2799-1/\$20.00

DOI: 10.1200/JCO.2008.19.6147

### ABSTRACT

#### Purpose

Many patients who present with early-stage colorectal cancer (International Union Against Cancer TNM stages I and II) are nevertheless at high risk of relapse. We hypothesized that intratumoral immune reaction could influence their prognosis.

#### Patients and Methods

The intratumoral immune reaction was investigated in 29 tumors by large-scale real-time polymerase chain reaction. Cytotoxic (CD8) and memory (CD45RO) T cells were quantified by immunohistochemical analyses of tissue microarrays from the center (CT) and the invasive margin (IM) of the 602 tumors from two independent cohorts. The results were correlated with tumor recurrence and patient survival.

#### Results

Patients with a strong infiltration of CD45RO<sup>+</sup> cells in the tumor exhibited an increased expression of T-helper 1 and cytotoxicity-related genes. Densities of CD45RO<sup>+</sup> and CD8<sup>+</sup> cells in tumor regions (CT/IM) classified the patients into four distinct prognostic groups based on the presence of high density of each marker in each tumor region. The four groups were associated with dramatic differences in disease-free, disease-specific, and overall survival (all  $P < .0001$ ). Five years after diagnosis, only 4.8% (95% CI, 0.6% to 8.8%) of patients with high densities of CD8<sup>+</sup> plus CD45RO<sup>+</sup> cells had tumor recurrence, and 86.2% (CI, 79.4% to 93.6%) survived. In contrast, the tumor recurred in 75% (95% CI, 17% to 92.5%) of patients with low densities of these cells, and only 27.5% (95% CI, 10.5% to 72%) survived (all  $P < .0001$ ). Multivariate analyses showed that the immune criteria had independent effects on the rates of complete remission and survival.

#### Conclusion

The combined analysis of CD8<sup>+</sup> plus CD45RO<sup>+</sup> cells in specific tumor regions could provide a useful criterion for the prediction of tumor recurrence and survival in patients with early-stage colorectal cancer.

*J Clin Oncol* 27. © 2009 by American Society of Clinical Oncology

### INTRODUCTION

Colorectal cancer is one of the most common malignancies for both men and women, with an estimated annual incidence of more than 1 million cases and approximately half a million deaths worldwide.<sup>1</sup> The frequency of localized colorectal cancers (stages I and II; International Union Against Cancer [UICC] TNM classification<sup>2</sup>) is approximately 40% and may indeed increase, as cancer societies are now recommending that the use of early detection procedures be started for low-risk adults at the age of 50 years.<sup>3</sup> Despite surgery with curative intent, the risk of disease recurrence for patients with stages I and II disease is high (approximately 20% to 30%). To subject all of these patients to postoperative chemo-

therapy may be inappropriate and costly.<sup>4</sup> Genetic and molecular tumor prognostic factors have been proposed to identify patients who may be at risk for recurrence. None of these have, however, been sufficiently informative for inclusion in clinical practice.<sup>5</sup> The identification of patients with a high-risk of disease recurrence therefore remains a major clinical issue.

Previous studies have suggested that the immune infiltrates in colorectal cancer are of clinical importance.<sup>6-11</sup> We showed that a high intratumoral memory T-cell density is correlated with a decreased incidence of tumor spread, not only to lymphovascular and perineural structures, but also to lymph nodes and distant organs.<sup>12</sup> We next demonstrated that the functional orientation, density, and location

of immune cells within colorectal tumors profoundly influences the clinical outcome for the patients, regardless of cancer stage.<sup>13</sup> Time to recurrence and overall survival times could be governed in large part by the state of the local adaptive immune reaction, even in early-stage carcinoma. Thus our previous results suggested that some of the patients with stages I and II colorectal cancer and a weak in situ immune reaction had indeed a poor prognosis.<sup>13,14</sup> We present herein a large study on 602 patients with stages I and II disease from two independent cohorts demonstrating the importance of the localized immune reaction in predicting recurrence and survival in patients with early-stage colorectal carcinoma.

## PATIENTS AND METHODS

### Study Populations

All patients with stages I and II colorectal cancer ( $n = 411$ ) who underwent a primary resection at the Laennec/Hôpital Européen Georges Pompidou (HEGP; Paris, France) between 1986 and 2004 were retrospectively reviewed and represented a prospective, continuous, unselected cohort of patients. Histopathologic and clinical findings were scored according to the TNM staging system of the UICC.<sup>2</sup> Two hundred twelve paraffin-embedded tumor samples (second cohort) from consecutive patients with stages I to II colorectal cancer between 1990 and 1991 were obtained from the federal state of Styria (Hospital of the Medical University of Graz [MedUG], Graz, Austria). The study was approved by the ethics committee of the MedUG.

### Real-Time Taqman Polymerase Chain Reaction Analysis

Twenty-nine frozen tumor specimens from patients with stages I and II colorectal cancer, with focused high- or low-memory T-cell density in the center and the invasive margin of the tumor, were extracted for RNA (see Appendix, online only). Gene expression was displayed using the Genesis program.<sup>15,16</sup>

### Tissue Microarray Construction

Permanent section hematoxylin and eosin (H&E)–stained slides were reviewed without knowledge of clinical characteristics or outcome. Each section H&E-stained slide was evaluated for the presence of the center of the tumor (CT) and the invasive margin (IM). For each tumor specimen, H&E-stained slides containing CT and IM regions were selected. For each region, two cores were taken from the areas containing the highest density of immune cells. For patients from Laennec/HEGP Hospital (France) and from MedUG (Austria), a manual tissue-array instrument (Beecher Instruments, Alphelys, Plaisir, France) and a fully robotized tissue-arrayer (ORIDIS Biomed, Graz, Austria) were used respectively (see Appendix, online only).

### Immunohistochemistry

For immunohistochemistry, tissue microarray (TMA) sections were incubated (60 minutes at room temperature) with monoclonal antibodies to CD45RO (OPD4), CD8 (4B11), and cytokeratin-8 (Neomarkers, Fremont, CA). Envision+ system and DAB-chromogen were applied (Dako, Copenhagen, Denmark) and counterstained with methylene blue. Slides were analyzed using an image analysis workstation (Spot Browser, Alphelys) whose reproducibility was previously validated.<sup>13</sup> Measurements were recorded as the number of positive cells per tissue surface unit (in square millimeters; see Appendix, online only).

### Statistical Analysis

We used the Wilcoxon–Mann-Whitney test to identify genes with significantly different levels of expression among patient groups (Fig 1). Kaplan-Meier curves were used to visualize differences of disease-free, disease-specific, and overall survival for patient groups stratified on CD45RO (CT/IM) and CD8 (CT/IM), and significance was assessed by univariate analysis with the use of the log-rank test. The “minimum  $P$  value” approach was used to assess the cutoff for the best separation of patients referring to their disease-free survival outcome.  $P$  values were corrected by applying the method proposed by Altman

et al<sup>17</sup> or using 100-fold cross-validations.<sup>18</sup> We used a multivariate Cox proportional hazards regression model to determine the hazard ratio of the “Immune Score” combination (CD45RO/CD8) together with clinical and histopathologic tumor markers. Hazard ratios were corrected as suggested by Holländer et al.<sup>19</sup> A  $P$  value less than .05 was considered statistically significant. All analyses were performed with the R (survival package; <http://www.r-project.org/>) and Statview (Abacus Systems, Berkeley, CA).

## RESULTS

### Immune Orientation and Memory T-Cell Infiltration in Stage I to II Colorectal Cancers

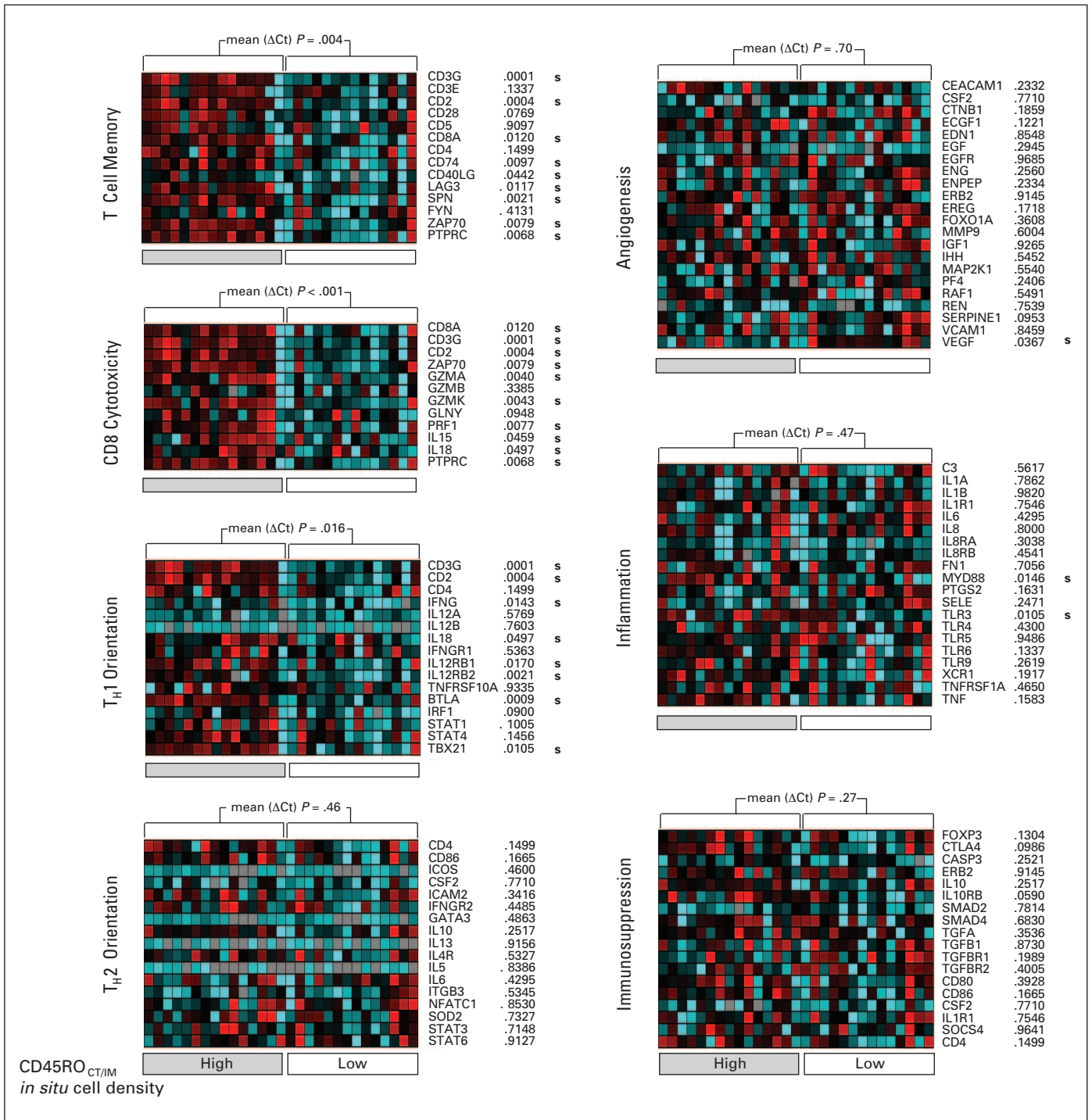
Memory T cells could be major players in the tumor microenvironment by influencing tumor spread and recurrence.<sup>12</sup> No information is however available concerning the immune environment favoring the emergence of memory T cells in colorectal cancer.

Gene expression levels related to T-cell memory, CD8 cytotoxicity, T helper 1 ( $T_{H1}$ ) and  $T_{H2}$  orientation, inflammation, immune suppression, and angiogenesis were assessed in 15 tumors with a high density of memory ( $CD45RO^{+}$ ) cells in the CT and the IM of the tumor ( $CD45RO_{CT/IM}^{Hi}$ ) and compared with the levels in 14 tumors with a low density of  $CD45RO^{+}$  cells in CT and IM ( $CD45RO_{CT/IM}^{Lo}$ ; Fig 1). As a control, we observed that the memory T-cell cluster was upregulated in  $CD45RO_{CT/IM}^{Hi}$  patients, whereas  $CD45RO_{CT/IM}^{Lo}$  patients presented with a significant downregulation of this cluster ( $P = .004$ ). A strong and coordinated in situ memory T-cell infiltration ( $CD45RO_{CT/IM}^{Hi}$  patients) was associated with the upregulation of the clusters referring to CD8 cytotoxicity and  $T_{H1}$  orientation ( $P < .001$  and  $P = .016$ , respectively, between patients groups; Fig 1). A heat map showed that 10 genes of the CD8 cytotoxicity cluster (eg, CD8, granzymes [*GZMA* and *GZMK*], perforin [*PERF1*]) and eight genes of the  $T_{H1}$  orientation cluster (eg, T-bet [*TBX21*], interferon gamma [*IFN $\gamma$* ], interleukin [*IL*] 12Rb 1 and 2, and IL-18) were significantly differentially expressed between patient groups (Fig 1). In contrast, the expression levels of the clusters associated with  $T_{H2}$  orientation, inflammation, immune suppression, and angiogenesis did not obviously influence  $CD45RO_{CT/IM}$  tumor infiltration (Fig 1). These data suggested that a high density of memory  $CD45RO^{+}$  cells in tumor regions (CT and IM) is found in a context of  $T_{H1}$  immune orientation and CD8 cytotoxic effectors. We therefore focused subsequent analyses on cytotoxic (CD8) and memory (CD45RO) T cells and hypothesized that each marker could provide a parameter of clinical outcome in patients with stages I to II disease.

### Immune Infiltration by Cytotoxic and Memory T Cells and Clinical Outcome

The 411 patients with stages I and II colorectal tumors are hereafter called the original cohort from the HEGP hospital. Among the clinical parameters of this cohort, depth of tumor invasion (pT stage) and bowel perforation at presentation were significantly associated with survival.

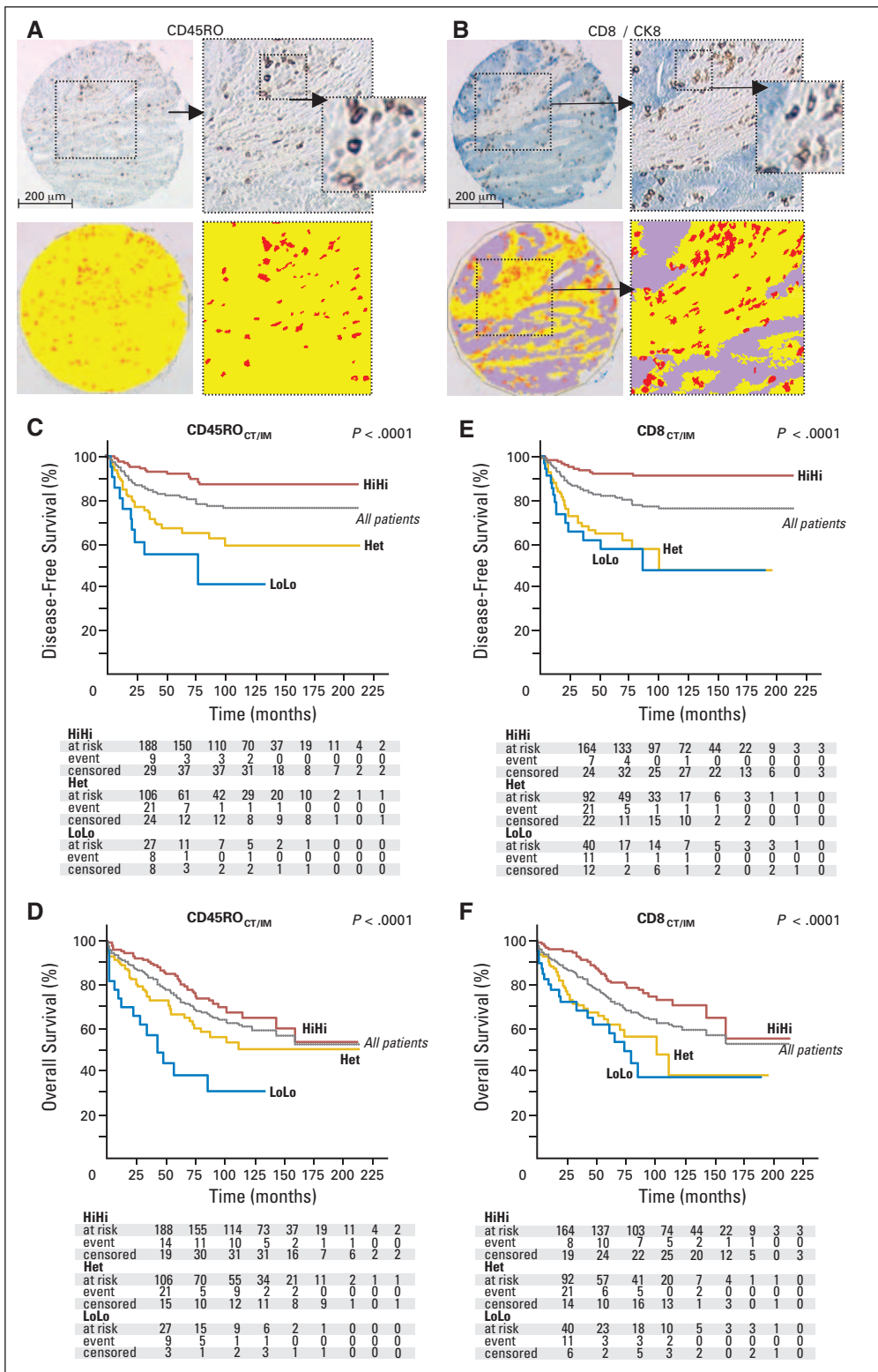
The densities of  $CD45RO^{+}$  and  $CD8^{+}$  immune infiltrates in the CT and in the IM were assessed for the 411 patients by immunohistochemical-based TMA analyses (Figs 2A and 2B) with image analysis software.<sup>13</sup> We showed, for CD45RO and CD8 markers, a statistically significant correlation between positive cell densities



**Fig 1.** Gene expression levels from 29 colorectal tumors with coordinated memory T-cell reaction determined by immunohistochemistry (15 CD45RO<sub>CT/IM</sub><sup>Hi</sup> v 14 CD45RO<sub>CT/IM</sub><sup>Lo</sup>) are shown. Gene expression levels were assessed by quantitative polymerase chain reaction and determined using threshold cycle (Ct) values normalized to 18S ( $\Delta Ct$ ). Heat-map representation of the genes clusters for different immune functions is shown. Genes are plotted from the minimal level of expression (blue) to the maximal level (red). The Mann-Whitney test was used to compare the expression levels of each gene between patient groups. Most of the genes of the clusters referring to CD8 cytotoxicity and T<sub>H</sub>1 orientation were significantly differentially expressed between patients groups. Of note, vascular endothelial growth factor, TLR3, and MYD88 were also correlated with CD45RO<sub>CT/IM</sub> infiltration.

in each tumor region (CT, IM) and patient outcome. Longer disease-free, disease-specific, and overall survival was observed among patients with tumors containing a high density of such cells than among patients whose tumors had low densities (log-rank  $P$  values ranging from  $P = .0087$  to  $P < .0001$ ).

For each immune marker, the combined analysis of CT plus IM regions (CT/IM) strongly discriminated patient groups. Compared with HiHi patients (hazard ratio [HR] = 1), HRs were 3.77 and 5.71 for CD45RO<sub>CT/IM</sub><sup>Het</sup> and CD8<sub>CT/IM</sub><sup>Het</sup> and 6.17 and 6.72 for CD45RO<sub>CT/IM</sub><sup>LoLo</sup> and CD8<sub>CT/IM</sub><sup>LoLo</sup>, respectively. Kaplan-Meier



**Fig 2.** (A, B) Upper part: representative examples of CD45RO and CD8 immunostaining of colorectal cancer tissue microarray. Immunostained cells (brown) and tumor cells (blue). For each marker, an enlargement of a spot and an enlargement of stained cells are illustrated. Lower part: digital image analyzed using the image software (SpotBrowser), with tissue represented in yellow and stained cells represented in red. The densities of CD45RO<sup>+</sup> and CD8<sup>+</sup> cells in each tumor region (center of tumor [CT] and invasive margin [IM]) were recorded as the number of positive cells per unit tissue surface area. (C, D) Kaplan-Meier curves illustrate the duration of disease-free survival and overall survival according to the densities of CD45RO<sup>+</sup> cells. (E, F) Kaplan-Meier curves according to the densities of CD8<sup>+</sup> cells in combined analysis of the tumor regions (CT/IM). For each marker, high densities in the CT and in the IM (HiHi; red line), low densities in the CT and in the IM (LoLo; blue line), and heterogeneous densities with CT<sup>Lo</sup> plus IM<sup>Hi</sup> or CT<sup>Hi</sup> plus IM<sup>Lo</sup> (Het; gold line) in patients with stages I and II colorectal cancer (log-rank statistical test,  $P < .001$  for all comparisons). Duration of disease-free survival and overall survival of the entire cohort of patients is also represented (gray line).

curves illustrated longer disease-free survival, disease-specific survival, and overall survival among patients with high densities of CD45RO<sup>+</sup> cells in both tumor regions (CD45RO<sub>CT/IM</sub><sup>Hi</sup>, all log-rank tests,  $P < .0001$ ; Figs 2C and 2D). A low density of CD45RO<sup>+</sup> cells in a single tumor region (heterogeneous pattern LoHi or HiLo;

CD45RO<sub>CT/IM</sub><sup>Het</sup>) markedly impaired the beneficial effect of the immune cell reaction at the tumor site. In univariate analysis, CD45RO<sub>CT/IM</sub> presented with a higher HR and a smaller  $P$  value as compared with histopathologic parameters. Similar results were obtained regarding the prognostic value of CD8<sup>+</sup> cells in combined

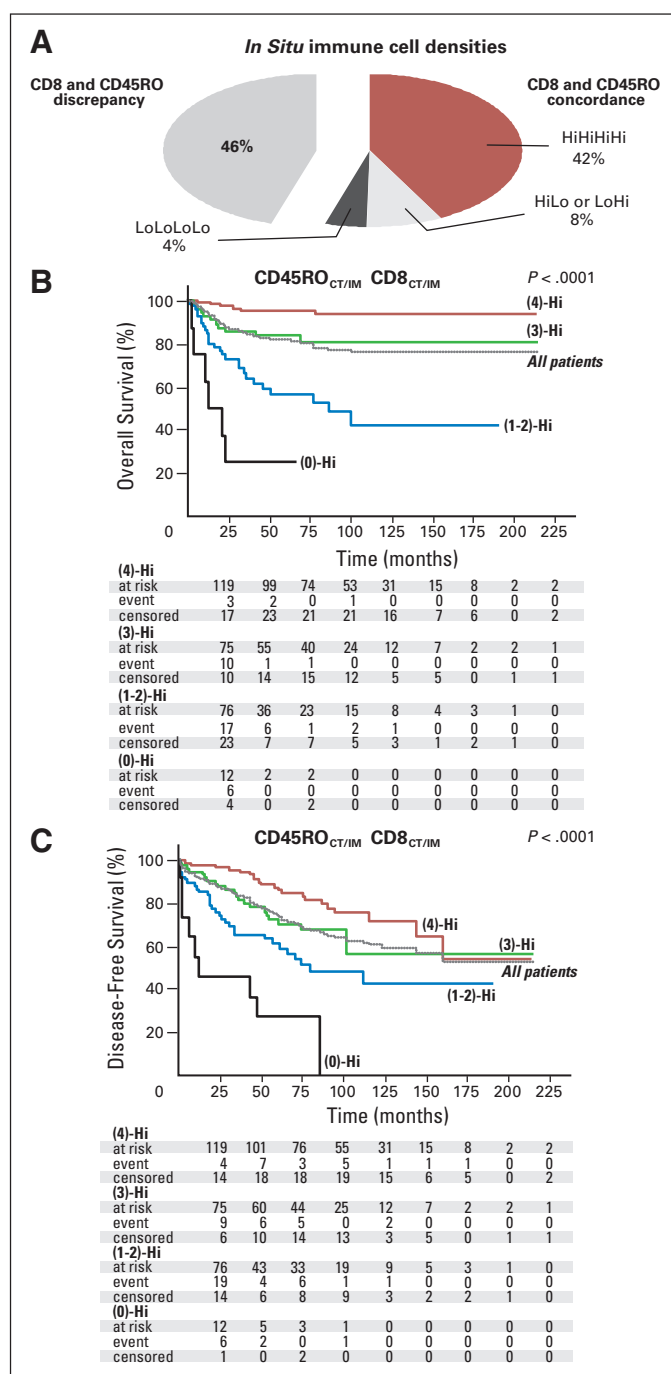
tumor regions (Figs 2E and 2F). We analyzed the influence of preoperative treatments on the densities and distributions (Hi/Lo) of CD8<sup>+</sup> and CD45RO<sup>+</sup> cells in tumor regions (CT and IM). No significant differences were found between patient groups (all  $P > .05$ ). To further validate the importance of CD45RO and CD8, the data were analyzed using 100 repetitions of two-fold cross-validation with and without stratifying on the grouping variable.<sup>18</sup> CD45RO<sub>CT/IM</sub> and CD8<sub>CT/IM</sub> remained highly significant parameters (all  $P < .0001$ ) associated with disease-free survival, disease-specific survival, and overall survival. This indicated that cytotoxic and memory T-cell localization and densities within the tumor strongly influenced prognosis of patients with stages I and II colorectal cancer.

**Combined Assessment of Cytotoxic and Memory Phenotypes to Predict Clinical Outcome**

Forty-two percent of the patients presented with a high infiltration of CD8<sup>+</sup> and CD45RO<sup>+</sup> cells in CT and IM regions of the tumor, whereas 4% had a low infiltration of these cells in both tumor regions (Fig 3A). Strikingly, 46% of the patients presented with discrepancies between the densities of the immune markers. We investigated whether a combined assessment of the memory and cytotoxic phenotypes could increase the accuracy of prediction of clinical outcome for the different patient groups.

As each marker in each tumor region positively influenced clinical outcome, we stratified the patients according to the total number of high densities (Hi) observed in the tumor (two markers assessed in two tumor regions, CT and IM). According to this criterion, 4%, 10%, 17%, 27%, and 42% of the patients presented with an immune score (IS) of (0)-Hi (1)-Hi, (2)-Hi, (3)-Hi, and (4)-Hi, respectively. Significant differences in disease-free, disease-specific, and overall survival times among patients groups were determined by univariate analysis ( $P < .0001$ ) and illustrated by Kaplan and Meier curves. Patients with high densities of CD45RO<sup>+</sup> and CD8<sup>+</sup> cells in both tumor regions ([4]-Hi) were at low risk, with 5-year disease-free and overall survival rates of 95.2% (CI, 91.2% to 99.4%) and 86.2% (CI, 79.4% to 93.6%), respectively. In contrast, patients with low densities of CD45RO<sup>+</sup> and CD8<sup>+</sup> cells in both tumor regions ([0]-Hi) were at higher risk, with 5-year disease-free and overall survival rates of 25% (CI, 7.5% to 83%; HR = 23.08; 95% CI, 7.6% to 70.06%) and 27.5% (CI, 10.5% to 72%; HR = 11.53; 95% CI, 5.01% to 26.52%), respectively. Patients with an IS of 1 or 2 high ([1]-Hi, [2]-Hi) experienced a similarly bad post-operative outcome. The cumulative 5-year disease-free and overall survival rates of the grouped patients ([1-2]-Hi) were 56.4% (HR = 10.18; 95% CI, 4.33% to 23.95%) and 61.6% (HR = 3.18; 95% CI, 1.84% to 5.51%), respectively. Finally, for patients with 3-Hi, the survival rate did not differ significantly from that observed in the entire cohort of patients with stages I and II disease (Figs 3B and 3C). Similar results were found for univariate analyses using categoric values for all of the parameters.

Additional analyses with 100 repetitions of two-fold cross-validation confirmed the highly significant correlation (all tests  $P < .0001$ ) between the IS and clinical outcome. We next found significant differences for patients with colon cancer and rectal cancer after separate stratification according to IS (all log-rank tests,  $P < .0001$ ). The IS, furthermore, remained a significant prognostic factor in the subgroups of patients with stage I and stage II colorectal cancers (all log-rank tests,  $P < .0001$ ). Using Cox multivariate regression analyses (Table 1), the IS and the bowel perforation at presenta-



**Fig 3.** (A) The pie chart illustrates the repartition of the patients according to the presence of concordance (right part) or discrepancy (left part) between CD8 and CD45RO<sup>+</sup> cell densities in combined tumor regions of stages I and II colorectal tumors. Fifty-four percent of the patients presented with similar high (4)-Hi, low (0)-Hi, or heterogeneous (2)-Hi immune infiltration for markers, whereas 46% presented with distinct patterns. (B, C) Kaplan-Meier curves for the duration of disease-free survival and overall survival according to a combined analysis of CD8<sup>+</sup> and CD45RO<sup>+</sup> densities in tumor regions (center of the tumor [CT] and invasive margin [IM]), in patients with stage I or II colorectal cancer. Patients are stratified according to an immune score ranging from 0 to 4, depending on the total number of high densities observed (two markers assessed in CT, two markers assessed in IM). For example, (4)-Hi refers to a tumor with high densities of CD8<sup>+</sup> and CD45RO<sup>+</sup> cells in CT and IM regions of the tumor (red line). (3)-Hi refers to tumors with three high densities (green line). (0)-Hi represents tumors with low densities of CD8<sup>+</sup> and CD45RO<sup>+</sup> cells in both tumor regions (black line). Patients with an immune score of 1 or 2 [(1)-Hi, (2)-Hi] experiencing similar outcome were therefore regrouped [(1-2)-Hi; blue line]. Log-rank statistical test,  $P < .001$  for all comparisons. Duration of disease-free survival and overall survival of the entire cohort of patients is also represented (gray line).

**Table 1.** Multivariate Cox Proportional Hazards Analysis of Patients With International Union Against Cancer TNM Stage I/II Colorectal Cancer

Variable	PHA Test*	Hazard Ratio	95% CI	P
<b>DFS</b>				
T stage	0.7817	1.17	0.81 to 1.68	.4100
Perforation†	0.8840	5.51	2.17 to 14.03	.0003‡
Immune score	0.7243	0.34	0.24 to 0.48	< .0001‡§
		0.35	0.25 to 0.49	
<b>OS</b>				
T stage	0.4476	1.03	0.81 to 1.33	.7900
Perforation†	0.3728	5.38	2.45 to 11.78	< .0001‡
Immune score	0.0229¶	0.54	0.43 to 0.69	< .0001‡
<b>DSS</b>				
T stage	0.8670	1.22	0.79 to 1.88	.3800
Perforation†	0.3559	11.05	4.05 to 30.2	< .0001‡
Immune score	0.4282	0.35	0.24 to 0.52	< .0001‡

NOTE. Immune score: CD8<sub>CT/IM</sub>CD45RO<sub>CT/IM</sub> (four groups: (4)-Hi, (3)-Hi, (2-1)-Hi, (0)-Hi). Similar results obtained using bootstrap.  
Abbreviations: PHA, proportional hazards assumption; DFS, disease-free survival; T, tumor; OS, overall survival; DSS, disease-specific survival.  
\*PHA test ( $P < .05$  violates the hazards assumption).  
†Bowel perforation presentation (yes, no).  
‡Significant.  
§Minimum  $P$  value cutoff with four groups: (4)-Hi, (3)-Hi, (2-1)-Hi, (0)-Hi.  
||Leave-one-out method, correction using  $C = 1 - (SE[\text{coef}]/\text{coef})^2 = 0.980$ .  
¶Heuristic shrinkage factor, corrected by Holländer et al.<sup>19</sup>  
¶¶The immune score violates the hazards assumption for OS and cannot be considered for OS, although it is highly significant.

tion were found to be significantly and independently associated with disease-free, disease-specific, and overall survival (IS: HR = 2.91, 2.82 and 1.85, respectively, all  $P < .0001$ ).

We validated the prognostic value of the IS in an independent cohort of 191 patients with stages I and II colorectal cancers from the Austrian Federal state of Styria. Five years after diagnosis, the incidence of patients without disease progression and overall survival was significantly higher in tumors exhibiting a high IS ([4]-Hi) than for patients with low densities of such cells ([0]-Hi or [1-2]-Hi), as illustrated by the Kaplan-Meier survival curves. The IS was significantly associated with disease-free, disease-specific, and overall survival (HR ranging from 1.35 to 2.10; log-rank  $P$  values ranging from .03 to .0017). Finally, to go further to a routine, easy-to-do methodology, we investigated in a limited series of 45 randomly selected patients whether a department of pathology using a routine methodology would reproduce the results we obtained with the TMA/automated microscopic immunoanalysis system. Immunostained slides of whole tumor sections from the same donor blocks used to build TMAs were analyzed by a pathologist and categorized into groups (Hi or Lo; optical assessment) according to the cutoff values we defined. Concordant results with regard to the prognostic value of the IS were obtained.

## DISCUSSION

Our study demonstrated that the extent of the immune reaction at the tumor site is correlated with the clinical outcome in two large independent cohorts of patients with localized stages I and II colorectal cancer. Evaluation of CD8<sup>+</sup> cytotoxic T cells and CD45RO<sup>+</sup> memory

T cells in the center and the invasive margin of the tumor could provide a means for predicting tumor recurrence and patient survival.

By combining immunohistochemical, genomic, and bioinformatic analyses, we showed that a high density of infiltrating T cells with memory phenotype in stages I and II tumors is developed in a context of T<sub>H</sub>1 immune orientation (T-bet, IFN $\gamma$ , IL12R $\beta$ 1, IL12R $\beta$ 2, IL-18) and CD8 cytotoxic effectors (CD8, perforin, and granzymes). Notably, one angiogenic marker (vascular endothelial growth factor) was inversely correlated with CD45RO<sub>CT/IM</sub> tumor infiltration, and two markers (TLR3 and MYD88) from the inflammation cluster were correlated with CD45RO<sub>CT/IM</sub> infiltration. Interestingly, TLR3 also enhances the dendritic cell-mediated production of IFN $\gamma$  by Ag-specific T lymphocytes, and MYD88 is also involved in the signaling cascade mediated by the IFN $\gamma$ -inducing cytokine IL-18. In contrast, the immune functions related to inflammation, Th2 polarization, and immune suppression did not obviously influence memory T-cell infiltration.

Memory cells have trafficking properties and can be maintained for many years in the body.<sup>20</sup> Micrometastases and occult tumor cells can be detected in blood, bone marrow, and lymph nodes even in early-stage colorectal tumors.<sup>21,22</sup> Systemic cytotoxic and memory T cells may therefore be exposed to disseminated tumor cells.<sup>23</sup> Analysis of in situ cytotoxic and memory T cells could provide information concerning the existing cytotoxic capacity and the ability of the immune system to respond faster on tumor cell re-exposure or maintain an equilibrium state with cancer.<sup>24</sup>

Ultimately, evaluation of CD8<sup>+</sup> cytotoxic T cells and CD45RO<sup>+</sup> memory T cells in tumor regions successfully identified patients at high risk of tumor recurrence and death. Long-term results for survival and relapse submitted to multiple statistical settings indicated a concordant prognostic value of the combined CD8 and CD45RO assessment. Internal cross-validations and investigations on an independent cohort of patients suggest the reliability of the results. According to our data, patients can be stratified into four prognostic categories with significantly different disease-free survival and overall survival. The behavior of patients with a low infiltrate of (0)-Hi or (1-2)-Hi drastically differed from that of patients with a high immune score of (4)-Hi, with corresponding HRs for disease-free survival of 23.1 or 10.2, respectively. This indicates that a stratification of the patients with the evaluation of CD8<sup>+</sup> and CD45RO<sup>+</sup> cells could be of interest in clinical practice. In multivariate analysis, the IS and bowel perforation presentation remained independent prognostic factors. Of note, perforation was only present in 3.4% of the patients, and this feature is decreasing with earlier detection of colorectal cancer in recent years. In contrast, high- and low-risk patients ([0]-Hi or [1-2]-Hi and [4]-Hi, respectively) represented 73% of cases of the series. Previous studies suggesting a protective role of immune infiltrates in colorectal tumors<sup>6-11</sup> have been contradicted by other reports that did not reach this conclusion.<sup>25</sup> Variable descriptions of the type and subjective grading of host lymphocytic reaction has favored non-reproducibility of previous findings. Our investigations combined analyses of the functional orientation, the precise density (number of cells per square millimeter), and the location of immune cell populations in distinct tumor territories.<sup>13</sup> Importantly, this methodology uses a routine immunohistochemical technique to detect two membrane receptors. We also showed that the assessment of CD8 and

CD45RO could be performed on whole slides with an optical evaluation of immune cell densities, with concordant results with regard to the prognostic value of the IS.

In summary, our study provides strong evidence for the usefulness of the combined evaluation of memory and cytotoxicity for the prediction of tumor recurrence and survival in patients with stages I and II colorectal cancer. To be included in clinical practice, a new biologic marker should have a strong and reproducible impact on outcome, independent of tumor invasion. The immune criteria we defined herein could address these conditions.

#### AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

#### AUTHOR CONTRIBUTIONS

**Conception and design:** Franck Pagès, Jérôme Galon

#### REFERENCES

1. Parkin DM, Bray F, Ferlay J, et al: Global cancer statistics, 2002. *CA Cancer J Clin* 55:74-108, 2005
2. Sobin L, Wittekind C: TNM classification of malignant tumors. New York, NY, Wiley-Liss, 2002
3. Jemal A, Siegel R, Ward E, et al: Cancer statistics, 2006. *CA Cancer J Clin* 56:106-130, 2006
4. Benson AB 3rd, Schrag D, Somerfield MR, et al: American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. *J Clin Oncol* 22:3408-3419, 2004
5. Locker GY, Hamilton S, Harris J, et al: ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol* 24:5313-5327, 2006
6. Bremers AJ, Andreola S, Leo E, et al: T cell responses in colorectal cancer patients: Evidence for class II HLA-restricted recognition of shared tumor-associated antigens. *Int J Cancer* 88:956-961, 2000
7. Guidoboni M, Gafa R, Viel A, et al: Microsatellite instability and high content of activated cytotoxic lymphocytes identify colon cancer patients with a favorable prognosis. *Am J Pathol* 159:297-304, 2001
8. Jass JR: Lymphocytic infiltration and survival in rectal cancer. *J Clin Pathol* 39:585-589, 1986
9. Naito Y, Saito K, Shiiba K, et al: CD8+ T cells infiltrated within cancer cell nests as a prognostic

factor in human colorectal cancer. *Cancer Res* 58:3491-3494, 1998

10. Ropponen KM, Eskelinen MJ, Lipponen PK, et al: Prognostic value of tumour-infiltrating lymphocytes (TILs) in colorectal cancer. *J Pathol* 182:318-324, 1997
11. Zlobec I, Baker K, Minoo P, et al: Node-negative colorectal cancer at high risk of distant metastasis identified by combined analysis of lymph node status, vascular invasion, and Raf-1 kinase inhibitor protein expression. *Clin Cancer Res* 14:143-148, 2008
12. Pagès F, Berger A, Camus M, et al: Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* 353:2654-2666, 2005
13. Galon J, Costes A, Sanchez-Cabo F, et al: Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 313:1960-1964, 2006
14. Galon J, Fridman WH, Pages F: The adaptive immunologic microenvironment in colorectal cancer: A novel perspective. *Cancer Res* 67:1883-1886, 2007
15. Galon J, Franchimont D, Hiroi N, et al: Gene profiling reveals unknown enhancing and suppressive actions of glucocorticoids on immune cells. *FASEB J* 16:61-71, 2002
16. Sturn A, Quackenbush J, Trajanoski Z: Genesis: Cluster analysis of microarray data. *Bioinformatics* 18:207-208, 2002
17. Altman DG, Lausen B, Sauerbrei W, et al: Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* 86:829-835, 1994
18. Faraggi D, Simon R: A simulation study of cross-validation for selecting an optimal cutpoint in univariate survival analysis. *Stat Med* 15:2203-2213, 1996
19. Holländer N, Sauerbrei W, Schumacher M: Confidence intervals for the effect of a prognostic factor after selection of an "optimal" cutpoint. *Stat Med* 23:1701-1713, 2004
20. Sallusto F, Geginat J, Lanzavecchia A: Central memory and effector memory T cell subsets: Function, generation, and maintenance. *Annu Rev Immunol* 22:745-763, 2004
21. Tsavellas G, Patel H, Allen-Mersh TG: Detection and clinical significance of occult tumour cells in colorectal cancer. *Br J Surg* 88:1307-1320, 2001
22. Zach O, Lutz D: Tumor cell detection in peripheral blood and bone marrow. *Curr Opin Oncol* 18:48-56, 2006
23. Finn OJ: Cancer immunology. *N Engl J Med* 358:2704-2715, 2008
24. Koebel CM, Vermi W, Swann JB, et al: Adaptive immunity maintains occult cancer in an equilibrium state. *Nature* 450:903-907, 2007
25. Nanni O, Volpi A, Frassinetti GL, et al: Role of biological markers in the clinical outcome of colon cancer. *Br J Cancer* 87:868-875, 2002

#### Acknowledgment

We thank Alphelys (Plaisir, France) for help with tissue microarray experiments and digital capture and analysis; and Ion Gresser, MD, PhD, for critically reading the manuscript and for unwavering support. This article is dedicated to the memory of Professor Paul-Henri Cugnenc from the Digestive Surgery department of the Hôpital Européen Georges Pompidou, Paris, France.

#### Appendix

#### Supplementary Methods

**Study populations.** All patients with stages I and II colorectal cancer (n = 411) who underwent a primary resection at the Laennec/Hôpital Européen Georges Pompidou (Paris, France) between 1986 and 2004 were retrospectively reviewed and represented a prospective, continuous,

unselected cohort of patients. Histopathologic and clinical findings were scored according to the TNM staging system of the International Union Against Cancer (Sobin L, Wittekind C: TNM classification of malignant tumors. New York, NY, Wiley-Liss, 2002). The mean duration of follow-up was 62 months (range, 0.2 to 215 months). The maximum time until tumor recurrence and death was 100 and 160 months, respectively. Follow-up data were available for all patients in the cohort. For all patients, postsurgery surveillance was carried-out at Laennec/Hôpital Européen Georges Pompidou and associated hospitals according to general practice for patients with colon cancer. Two hundred twelve paraffin-embedded tumor samples (second cohort) from consecutive patients with stages I and II colorectal cancer between 1990 and 1991 were obtained from the federal state of Styria (Hospital of the Medical University of Graz, Graz, Austria). The study was approved by the ethics committee of the Medical University of Graz. The mean duration of follow-up was 101 months (range, 0.5 to 200 months). The maximum time until progression and death was 181 and 180 months, respectively. Twenty-one patients for whom follow-up data were unavailable were excluded from the analysis.

**Real-time Taqman polymerase chain reaction analysis.** Frozen tumor samples ( $n = 57$ ) of clinically localized colorectal cancer (International Union Against Cancer TNM stages I and II) from 1996 to 2004, with sufficient RNA quality and quantity for gene expression analysis, were selected. These tumors were assessed for CD45RO<sup>+</sup> cell densities in the center and the invasive margin by immunohistochemical analysis of tissue microarrays. Analyses of the data setting the cutoff at the median of the data sets showed that 15 patients presented with a high density of CD45RO<sup>+</sup> cells in both tumor regions (HiHi) versus 14 patients with a low density of CD45RO<sup>+</sup> cells in both regions (LoLo). The remaining patients presented with a heterogeneous pattern (LoHi or HiLo; CD45RO<sub>CT/IM</sub><sup>Het</sup>). All the 29 tumors with a coordinated CD45RO<sup>+</sup> immune infiltration (HiHi or LoLo) were assessed for gene expression analysis using Low-Density-Array (Applied Biosystems, Foster City, CA). Tissue samples were snap-frozen within 15 minutes after surgery and stored in liquid nitrogen. Twenty-nine frozen tumor specimens from patients with stages I and II colorectal cancer with focused high- or low-memory T-cell density in the center and the invasive margin of the tumor were extracted for RNA with the RNeasy isolation-kit (Qiagen, Valencia, CA). The integrity and the quantity of the RNA were evaluated on a bioanalyzer-2100 (Agilent Technologies, Palo Alto, CA). Immune-related gene expression was assessed by quantitative real-time TaqMan polymerase chain reaction (Low-Density-Arrays) and the 7900 robotic polymerase chain reaction system (Applied Biosystems). Data were analyzed using the SDS Software v2.2 (Applied Biosystems). Gene expression levels were determined using threshold cycle (Ct) values normalized to 18S ( $\Delta$ Ct). Results were displayed using the Genesis program ([www.genome.tugraz.at](http://www.genome.tugraz.at)).

**Tissue microarray construction and image analysis.** Tissue cylinders (diameter of 0.6 mm for the center of the tumor [CT] and 1 mm for the invasive margin [IM]) were punched from the targeted tumor areas of each donor tissue block and deposited into a recipient block using a tissue arraying instrument (Beecher Instruments, Alphelys, Plaisir, France). For patients from MedUG (Austria), a fully robotized tissue-arrayer was used (ORIDIS Biomed, Graz, Austria) with cores of 0.6 mm in diameter. After construction of the array block, the recipient blocks were subsequently cut into 5-mm sections on silanated glass slides for hematoxylin and eosin and immunohistochemical stainings. The hematoxylin and eosin staining was performed for histologic verification of the adequacy of the arrayed tumor tissues.

Slides were analyzed using an image analysis workstation (Spot Browser, Alphelys) whose reproducibility was previously validated (Galon J, Costes A, Sanchez-Cabo F, et al: *Science* 313:1960-1964, 2006). This system was used in preference to manual scoring because the capacity of observers to reproducibly score the number of positive cells on a continuous basis is limited. Validation studies were performed between optical and automatic cell counts. All comparisons showed a very high level of reproducibility achieved by the automated analysis workstation (Spot Browser).

**Immunohistochemistry.** Sections were incubated (60 minutes at room temperature) with monoclonal antibodies to CD45RO (OPD4), CD8 (4B11), and cytokeratin-8 (Neomarkers, Fremont, CA). Envision+ system and DAB-chromogen were applied (Dako, Copenhagen, Denmark). Tissue sections were counterstained with methylene blue. Irrelevant isotype-matched mouse monoclonal antibodies were used as negative controls. Slides were analyzed using an image analysis workstation (Spot Browser) whose reproducibility was previously validated (Galon J, Costes A, Sanchez-Cabo F, et al: *Science* 313:1960-1964, 2006). Polychromatic high-resolution spot images (740 × 540 pixels, 1.181  $\mu$ m/pixel resolution) were obtained (magnification ×100). Measurements were recorded as the number of positive cells per tissue surface unit (square millimeters). For each duplicate, the mean density was used for statistical analysis.

**Statistical analysis.** We used the Wilcoxon–Mann-Whitney test to identify genes with significantly different levels of expression among patient groups with high and low densities of CD45RO<sup>+</sup> (CT/IM) cells (Fig 1). Kaplan–Meier curves were used to visualize differences of disease-free, disease-specific, and overall survival for patient groups stratified on CD45RO (CT/IM) and CD8 (CT/IM) alone and in combination. The significance of disease-free, disease-specific, and overall survival of patients with early-stage colorectal cancer for various clinical characteristics and marker combinations was assessed by univariate analysis with the use of the log-rank test. The “minimum *P* value” approach was used to assess the cutoff for the best separation of patients referring to their disease-free survival outcome. *P* values were corrected applying the method proposed by Altman et al (Altman DG, Lausen B, Sauerbrei W, et al: *J Natl Cancer Inst* 86:829-835, 1994) or using 100-fold cross-validations (Faraggi D, Simon R: *Stat Med* 15:2203-2213, 1996). The cutoff values for CD8 and CD45RO cell densities were 83 and 110 cells/mm<sup>2</sup> in the CT and 102 and 221 cells/mm<sup>2</sup> in the IM, respectively. We used a multivariate Cox proportional hazards model to determine the hazard ratio of the “Immune Score” combination (CD45RO/CD8), together with clinical and histopathologic tumor markers. Hazard ratios were corrected as suggested by Holländer et al (Holländer N, Sauerbrei W, Schumacher M: *Stat Med* 23:1701-1713, 2004). A *P* value less than .05 was considered statistically significant. All analyses were performed with the R (survival package; <http://www.r-project.org/>) and Statview (Abacus Systems, Berkeley, CA) statistical programs.



## Glossary Terms

**CD8:** The surface antigen that characterizes CD8<sup>+</sup> T lymphocytes, CD8 is associated with the T-cell receptor and major histocompatibility complex, necessary for antigen recognition.

**CD45RO:** An isoform of CD45 that does not contain the coding sequence of exons A, B, and C. It is present on subsets of T and B cells, monocytes, and macrophages.

**Immunohistochemistry:** The application of antigen-antibody interactions to histochemical techniques. Typically, a tissue section is mounted on a slide and is incubated with antibodies (polyclonal or monoclonal) specific to the antigen (primary reaction). The antigen-antibody signal is then amplified using a second antibody conjugated to a complex of peroxidase-antiperoxidase (PAP), avidin-biotin-peroxidase (ABC) or avidin-biotin alkaline phosphatase. In the presence of substrate and chromogen, the enzyme forms a colored deposit at the sites of antibody-antigen binding. Immunofluorescence is an alternate approach to visualize antigens. In this technique, the primary

antigen-antibody signal is amplified using a second antibody conjugated to a fluorochrome. On UV light absorption, the fluorochrome emits its own light at a longer wavelength (fluorescence), thus allowing localization of antibody-antigen complexes.

**Prognostic factor:** A measurable patient characteristic that is associated with the subsequent course of disease (whether or not therapy is administered). The identification of a prognostic factor does not necessarily imply a cause-and-effect relationship. However, within a suitable outcome model, the measurement of a prognostic factor contributes to an estimate of an outcome probability (eg, the probability of disease-free survival within a given time interval).

**Cox proportional hazards regression model:** The Cox proportional hazards regression model is a statistical model for regression analysis of censored survival data. It examines the relationship of censored survival distribution to one or more covariates. It produces a baseline survival curve, covariate coefficient estimates with their standard errors, risk ratios, 95% CIs, and significance levels.