Haematology and Blood Transfusion Vol. 29
Modern Trends in Human Leukemia VI
Edited by Neth, Gallo, Greaves, Janka

Oncogenes of Avian Acute Leukemia Viruses are Subsets of Normal Cellular Genes

T. S. Papas[1], N. C. Kan[1], D. K. Watson[1], J. A. Lautenberger[1], C. Flordellis[1], K. P. Samuel[1], U. G. Rovigatti[1], M. Psallidopoulos[1], R. Ascione[1], and P. H. Duesberg[2]

A. Introduction

Avian acute leukemia viruses are a group of defective retroviruses which cause neoplasia in animals after short periods of latency. Transformation induced by the acute leukemia viruses usually manifest the direct expression of one or two viral *onc* genes that include characteristic elements transduced from cellular genes termed proto-oncogenes. Table 1 demonstrates the oncogenic properties of these viruses. Of particular interest to our laboratory is the MC29 subgroup of acute transforming viruses with *onc*-specific sequence termed *myc* as well as the avian myeloblastosis virus (AMV) subgroup with an *onc*-specific sequence termed *myb*. There are dramatic biologic differences between these two subgroups of viruses. As illustrated in Table 1, the MC29 subgroup possesses a broad spectrum of oncogenicity in vivo, causing leukemia and solid tumors in animals, and the ability to transform a variety of cells such as fibroblasts and hematopoietic cells in vitro. By contrast, the AMV subgroup has a much narrower spectrum of oncogenicity in vivo, causing leukemia, but not solid tumors, and possesses the ability to transform only hematopoietic cells in vitro.

Significant biologic differences exist among the viruses of the MC29 subgroup. To understand these biologic differences, we have cloned the integrated proviral genomes of the MC29 and MH2 viruses and analyzed their genetic structures. The *onc* gene of MC29 is a genetic hybrid (*Δgag-myc*), that consists of an element derived from the retroviral *gag* gene linked to the two 3′ exons of the cellular proto-*myc* gene. Thus in MC29-infected cells, the transforming *onc* gene product is expressed as a hybrid protein that includes both *gag* and *myc* sequences. A unique characteristic of the MH2 virus that has been described in our laboratory is that this virus contains two different *onc* genes, each of which is different from the *Δgag-myc* gene of MC29. One MH2 *onc*-gene has the genetic structure *Δgag-mht*, the other exists in two exons, a small 5′ exon of 6 *gag* codons and a major 3′ *myc* exon that is almost colinear with the *myc* sequence of MC29 [1, 2], Finally, our studies on the proto-*myc* and proto-*mht* genes have enabled us to reconstruct how the *myc* and *mht* sequences of MC29 and MH2 were transduced from proto-*myc* and proto-*mht* by the retroviruses that must have generated MC29 and MH2.

B. Results and Discussion

I. Sequence Analysis of MC29 and Chicken and Human Proto-*myc* Oncogenes

We have cloned [3] and sequenced the *Δgag-myc* gene of MC29 as well as the chicken proto-*myc* gene, and the cor-

1 National Cancer Institute-FCRF, National Institutes of Health, Frederick, MD 21701
2 Department of Molecular Biology, University of California, Berkeley, California 94740, USA

Table 1. Oncogenic properties of acute leukemia virus

Virus Strain	Neoplastic growth induced in vivo			Cell Types Transformed in vitro	Viral *onc* sequences
	Sarcoma (Fibrosarcoma = F) (Hepatocytoma = H)	*Carcinoma* (Renal adenocarcinoma = RC) (Carcinoma = C)	*Acute leukemia* (Myelocytomatosis = M) (Erythroblastosis = E) (Myeloblastosis = My)	(Fibroblastic = f) (Epitheloid = ep) (Myeloid = m) (Erythroid = e)	
MC29 subgroup					
MC29	F, H	C, RC	M, E	f, ep, m, e	*myc*
MH2	F, H	C, RC	M, E	f, ep, m, e	*myc, mht*
OK10	F, H	C, RC	M, E	f, ep, m, e	*myc*
CMII	F, H	C, RC	M, E	f, ep, m, e	*myc*
AMV subgroup					
AMV			My	m	*myb*
E26			My, E	m, e	*myb, ets*
AEV subgroup					
AEV	F	C	E, M	f, e	*erb* A, *erb* B

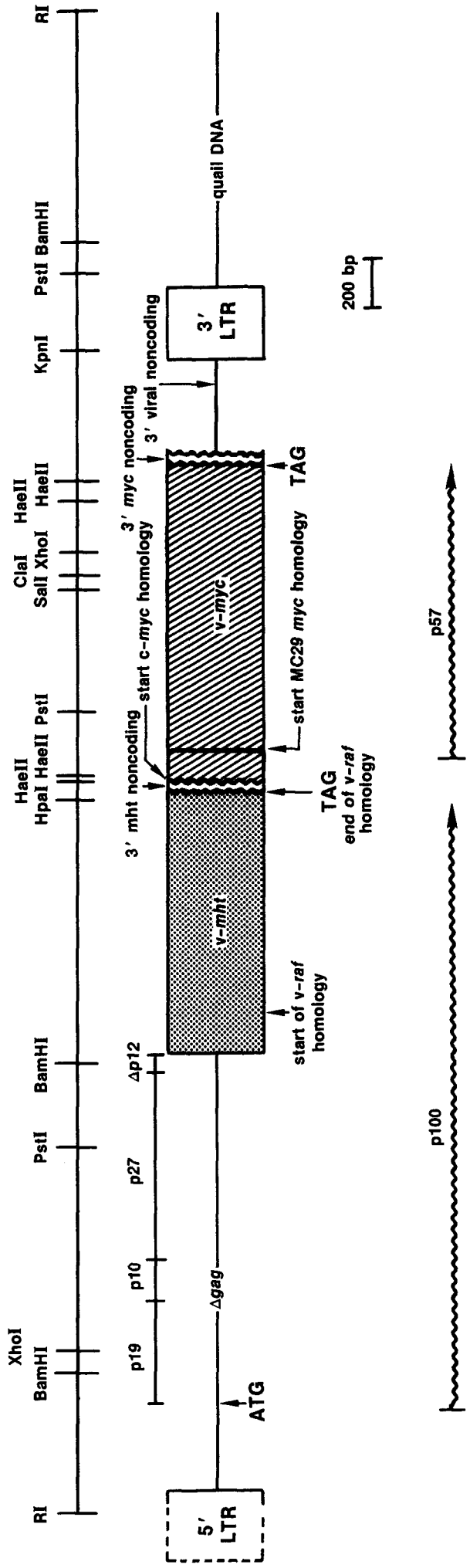
responding human proto-*myc* gene [4–6]. We have then aligned the sequences of the viral and cellular *myc*-related genes on the basis of their nucleotide homology, to identify similarities and dissimilarities, and have reached the conclusions presented elsewhere [7].

II. Molecular Structure of MH2

We have molecularly cloned the MH2 provirus, and determined its exact genetic structure by sequence analysis shown in Fig. 1 [2]. As expected, the virus contained a *gag*-derived sequence as the 5′ end of the viral genome and a *myc*-sequence as its 3′ end. Unexpectedly, this analysis revealed an MH2-specific sequence of 1.2 kilobases, termed *mht,* which is unrelated to the *myc* sequence and which maps between *Δgag* and *myc* [4, 5]. The nucleotide sequence of the MH2 viral genome indicates that the *gag* region and the *mht* gene form an open reading frame starting at the known AUG codon of the *gag* gene and terminating at a TAG stop codon near the 3′ end of *mht.* This open reading frame contains 894 amino acids capable of encoding a protein of about 100 kilodaltons. This prediction is in accord with the size of the p100 *gag-mht* protein observed in MH2-transformed cells [8].

III. Sequence Homology Between *mht* and *raf*

When the *mht* sequence of MH2 was compared with the *onc*-specific *raf* sequence of mouse sarcoma virus (MSV) 3611, a striking homology extending from the 5′ end of *mht* sequence for over 969 nucleotides was observed (Fig. 1) [9–10]. At the nucleotide level, the homology is 80%. Most nucleotide changes are third-base substitutions that result in no amino acid changes. The *raf* sequence differs from the *mht* sequence by having one inserted codon (proline) at position 1170 and by 19 amino acid substitutions. The absence of proline at position 1170 in the *mht* sequence is confirmed by MH2-specific oligonucleotide 6a [1]. Thus, the homology is 94% at the deduced amino acid level. This region of homology is flanked in both viruses by MH2- and



MSV-specific sequences with essentially no homology between the two viruses. At the 5′ end, the homology begins 174 base pairs 3′ to the *gag-mht* junction in MH2. Thus, the first 57 amino acids of *mht* preceding the start of homology with *raf* are MH2 specific, whereas *raf* is a completely colinear subset of *mht* [10]. At the 3′ end, the two sequences share a common termination codon at position 1186 and diverge beyond this position. Following the termination codon, there are 12 and 175 presumably cell-derived noncoding nucleotides in *mht* and *raf*, respectively.

IV. Relation of *mht* to Cellular proto-*mht* Genes

We have recently shown that the *mht* sequence, like the specific sequences of other retroviral *onc* genes, has a cellular counterpart in the chicken chromosome [7] as well as in the chromosomes of mice and humans (data not shown). Because a 3.8 kilobases cellular *mht*-related mRNA was detected by RNA blot hybridization in normal chicken cells [7], the chicken proto-*mht* gene may contain a coding sequence of about 3.8 kilobases. Therefore, the 1.3 kilobases *mht* sequence of MH2 probably represents only a subset of the proto-*mht* gene. The fact that the *mht* sequence terminates with several translation stop codons, suggests that it represents the 3′ end of the proto-*mht* gene. The same is true for the relation between the viral *myc* and proto-*myc* genes [7, 11]. Thus, the *Δgag-mht* gene and the cellular proto-*mht* gene are coterminal, but not isogenic. This is consistent with the

Fig. 1. Genetic structure of MH2 proviral genome. A restriction enzyme map of the 6.5 kilobases *Eco*RI-resistant quail DNA fragment that includes the MH2 provirus [1] is shown above the genetic map of MH2 proviral genome. *Broken lines* represent sequences not present in the 6.5 kilobases *Eco*RI fragment; *boxes* between the 5′ and 3′ LTR represent cell-derived sequences; *full lines* between the two LTR represent viral sequences; p100 and p57 represent the *Δgag-mht* hybrid protein and the *myc*-containing protein in MH2-transformed cells, respectively

transduced by avian acute leukemia viruses

Cellular gene	Transcript (kilobases)	Virus	Segment transduced (kilobases)
myc	2.5	MC29	1.4
		MH2	1.5
mht	3.8	MH2	1.2
myb	4.0	AMV	1.2
		E26	0.8
ets	7.4	E26	1.5

notion that the *onc* genes of most retroviruses are subsets of their cellular prototypes. By the same analogy, our parallel observation with AMV and MV29 viruses compared with their proto-*onc* transcripts seem to support this overall scheme (Table 2). The most dramatic example is the segment of the *ets* gene transduced by E26 virus which is four times smaller than the proto-*ets* transcript identified in normal chicken embryo fibroblasts (Table 2) [12, 13].

References

1. Kan N, Flordellis C, Garon C, Duesberg P, Papas T (1983) Proc Natl Acad Sci USA 80:6566
2. Kan N, Flordellis C, Mark G, Duesberg P, Papas T (1984) Proc Natl Acad Sci USA 81:3000
3. Lautenberger J, Schultz R, Garon C, Tsichlis P, Papas T (1981) Proc Natl Acad Sci USA 78:1518
4. Reddy E, Reynolds K, Watson D, Schulz R, Lautenberger J, Papas T (1983) Proc Natl Acad Sci USA 80:2500
5. Watson D, Reddy E, Duesberg P, Papas T (1983) Proc Natl Acad Sci USA 80:2146
6. Watson D, Psallidopoulos M, Samuel D, Dalla Favera R, Papas T (1983) Proc Natl Acad Sci USA 80:3642
7. Papas T, Kan N, Watson D, Flordellis C, Psallidopoulos M, Lautenberger J, Samuel K, Duesberg P (1984) In: Levin A, Topp W, Vande Woude GF, Watson JD (eds) Cancer cells. Cold Spring Harbor, New York, pp 153–160
8. Hu S, Moscovici C, Vogt P (1978) Virology 89:162
9. Rapp U, Godsborough M, Mark G, Bonner T, Groffen J, Reynolds F, Stephenson J (1983) Proc Natl Acad Sci USA 80:4218
10. Kan N, Flordellis C, Mark G, Duesberg P, Papas T (1984) Science 233:813
11. Watson D, Reddy E, Duesberg P, Papas T (1983) Proc Natl Acad Sci USA 80:2146
12. Leprince D, Genonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C, Stehelin D (1983) Nature 306:395
13. Nunn M, Seeburg P, Moscovici C, Duesberg P (1983) Nature 306:391